UNIVERSITY OF CANTERBURY
COLLEGE OF SCIENCE
WATERWAYS CENTRE FOR FRESHWATER MANAGEMENT
WATR 690A – MASTER OF WATER RESOURCE MANAGEMENT THESIS

QUANTIFICATION OF THE PROBABLE ENVIRONMENTAL EFFECTS OF THE HINDS MANAGED AQUIFER RECHARGE TRIAL USING MATHEMATICAL MODELLING AND ADVANCED UNCERTAINTY TECHNIQUES

Submitted by

Kim Patrick Durney

97191197

Supervisors

Dr Leanne Morgan (Primary Supervisor)

Mr David Scott (Co-Supervisor)

Dr Theo Sarris (Co-Supervisor)

Submission Date: 28/08/2019

Acknowledgements

Firstly, I would like to thank my primary supervisor, Dr Leanne Morgan. Not only were you instrumental in bringing this research to fruition, you always took the time to read and provide feedback on my work as it progressed.

I would also like to thank David Scott and Theo Sarris at ESR, your patience, supervision and advice as I developed the models and approaches used in this research have been critical to its completion.

Special thanks to: Aqualinc Research Ltd for their advice and data around Canterbury soils and infiltration; William Dench at Mayfield Hinds Valetta Irrigation Ltd for provision of data and keeping me informed as to the trials progress; Environment Canterbury for data provision and use of their Groundwater Vistas software licence; Matt Knowling for helping formulate the research questions; and Hisham Zarour for encouraging me to undertake further study.

I would also like to thank Bob Bower and Brett Sinclair at Wallbridge Gilbert Aztec; if I had not worked alongside you both, I would never have undertaken this research.

Finally, and certainly, not least, I would like to give my deepest gratitude to my family and principally my wife Jade for your support over the last two years. You not only singlehandedly kept our household functioning, but me too.

Abstract

Internationally, Managed Aquifer Recharge (MAR) has gained recognition as a mechanism to address environmental degradation. Investigations into its effectiveness in the New Zealand setting are ongoing and started with a five-year trial near Hinds, mid-Canterbury. The Hinds MAR trial aimed to raise groundwater levels, improve lowland streamflow and improve both ground and surface water quality. This research used advanced numerical modelling techniques such as null space Monte Carlo to assess the probable effectiveness of the Hinds MAR trial. Use of numerical uncertainty analysis to understand the probable effects of MAR programmes is both recommended and noted as a gap in international literature. Secondly, this research investigates the usefulness of various modelling approaches for quantifying the effects of MAR on the receiving environment.

The water resources of the Hinds Plains, mid-Canterbury, have been degraded by decades of high-intensity agriculture. The effects are seen as lowered groundwater levels, reduced spring-fed stream flows and high nitrate-nitrogen concentrations in both surface and groundwater. The Hinds MAR trial sought to introduce 500 l/s recharge into the groundwater system. The original evaluations of the trials probable effectiveness are based on this rate of recharge. However, due to in-situ conditions, the design recharge rate was never realised. Using the model code MODFLOW-NWT and advanced uncertainty analysis, this research investigated the probable response to the long-term achievable recharge rate of approximately 110 l/s after the trial completion.

This research investigated the effectiveness of a range of modelling approaches, from simple analytical and homogenous parameter numerical models to highly parameterised, spatially variable numerical models. It ultimately settles on highly parameterised numerical modelling as the most effective approach to assess the effects of the MAR trial. Like previous international studies, this research demonstrates the importance of calibrating numerical models to both quantity and quality, especially if useful results are to be obtained for water quality outcomes. For instance, initial uncertainty analysis used a model that was only calibrated to quantity, and despite a suite of 100 calibration constrained simulations it was not possible to reproduce observed water quality changes in response to the trial. Failure to capture the hydraulic conductivity field that represented the observed water quality changes was possibly caused by the model calibration being biased towards a local calibration optima, something highlighted in international literature as a possible unwanted outcome with calibration constrained uncertainty analysis. Recalibration of the flow model with the inclusion of concentration targets as calibration criteria produced more reasonable approximations of the observed and expected water quality changes.

The final suite of null space Monte-Carlo simulations suggest the Hinds MAR trial will successfully raise groundwater levels across a large area and increase stream flows. Further, the trial will improve water quality in groundwater, though it will probably not influence surface water quality. Transport modelling suggests water quality improvements can be expected for several kilometres down-gradient of the trial site, though they are unlikely to propagate as far as the lowland streams.

In terms of the appropriateness of the various modelling techniques investigated, analytical modelling is likely sufficient to estimate the mounding effects in the immediate vicinity of the trial. However, once the area of interest extends beyond the immediate trial site, numerical modelling should be applied. Water quality observations should be included in the calibration targets, and flow and transport calibrated simultaneously. Uncertainty analysis was useful for providing confidence in modelled outcomes and should be employed if the risks (whether financial or environmental) associated with MAR programmes need to be considered.

Contents

List of Figures

List of Tables

1 Introduction

Managed aquifer recharge (MAR) refers to the intentional augmentation of aquifer systems. Increasingly, MAR is seen as an effective approach for the sustainable management of both surface and groundwaters. MAR has been adopted internationally for more than six decades (Dillon et al. 2018). The importance of MAR has been recognised by the UNESCO organisation, the International Groundwater Resources Assessment Centre (IGRAC), which describes MAR as a very important tool for adaptation to climate change (IGRAC 2019). MAR is an emerging technology in New Zealand. Despite wide international adoption, New Zealand has been slow to uptake the approach. However, locally, interest has recently increased with the start of a five-year MAR study in 2016 in the Hinds catchment of mid-Canterbury. To date, as with the Hinds trial, most MAR programmes in New Zealand have been pilot studies, with the effectiveness of the technology in New Zealand still viewed with scepticism.

MAR is a catch-all term for several technologies that range from infiltration basins, to direct injection of water deep within confined aquifer systems. While there are many reasons for the adoption of MAR programmes, the one factor linking all these technologies is the intentional recharge of aquifer systems (Dillon et al. 2009).

The Hinds MAR trial sought to recharge the declining groundwater resource of the region by diverting alpine sourced Rangitata River water into a mid-catchment infiltration basin (Figure 1-1).

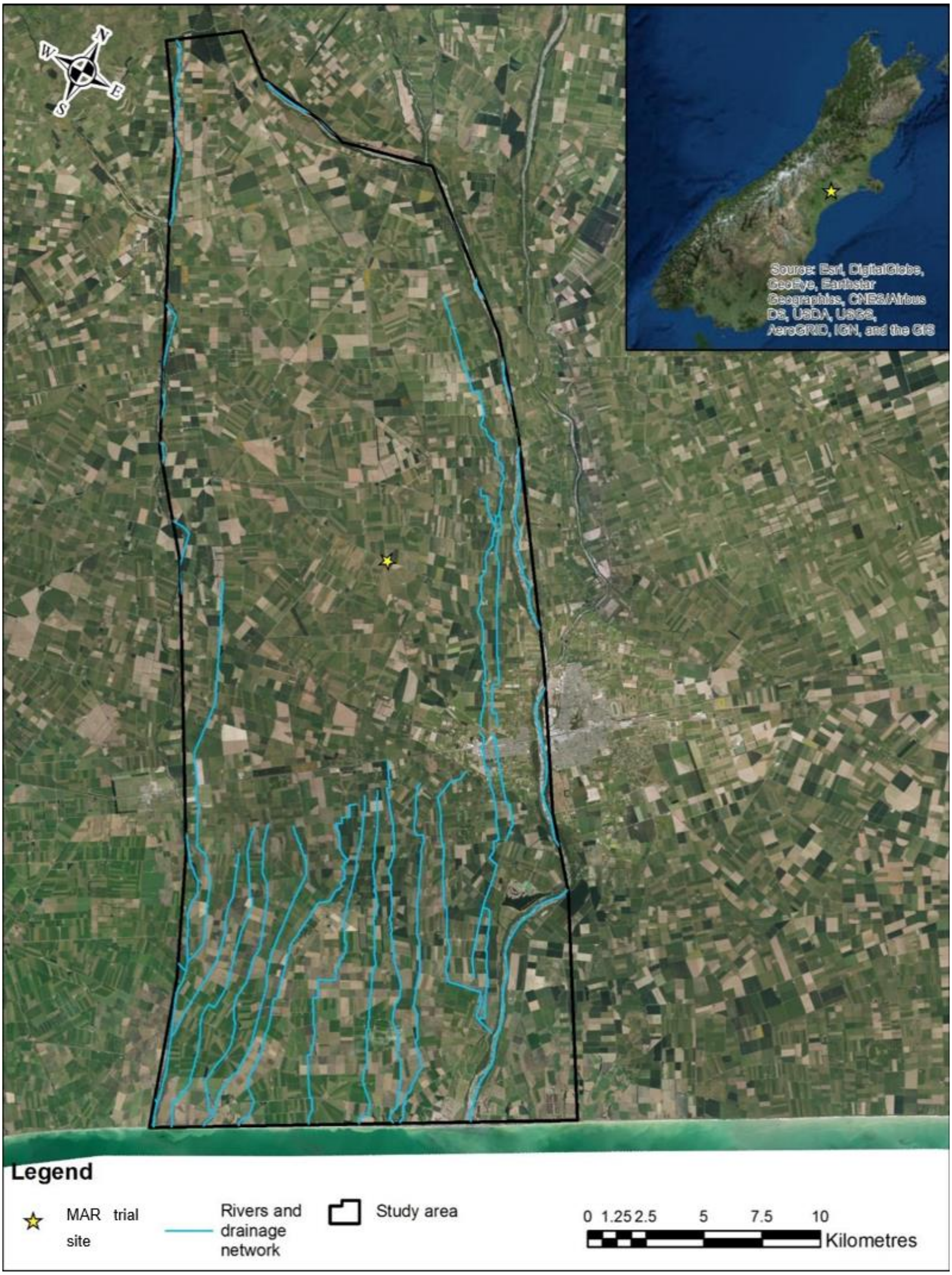

Figure 1-1 Hinds MAR trial site.

Specifically, the Hinds MAR trial aimed to (Golder 2017):

1. raise groundwater levels,
2. improve water quality in a highly polluted aquifer system and
3. increase streamflow in down-gradient spring-fed streams.

The Hinds MAR trial used a combination of an infiltration/spreading basin and dry wells[1] to infiltrate Rangitata River water into the unconfined aquifer beneath the trial site. Initial estimates of the recharge rate that could be sustained long-term from the infiltration basin were based on two pre-trial, field infiltration rate tests and deterministic[2] 2D unsaturated zone modelling (Golder 2015). Golder's infiltration modelling suggested that recharge rates between 0.5 m^3/s and 1.5 m^3/s could be sustained (Golder 2015). Based on these estimates of infiltration potential, Environment Canterbury (ECAN) then used numerical modelling to predict the extent of mounding and water quality changes resulting from the trial. The pre-trial 3D numerical modelling by ECAN used an existing model of the Hinds region that had been developed using the DHI software MIKE SHE (Durney et al. 2014). Based on this modelling and the pretrial testing, it was concluded that the water quantity goals of the trial were achievable. However, during the trial, the recharge rates were an order of magnitude lower than predicted (between 0.09 m^3/s and 0.14 m^3/s) (Golder 2017).

Revised 3D modelling was conducted using MODFLOW NWT, using the first six to nine months of trial data (Golder 2017). This MODFLOW modelling suggested that goals 1 and 2 of the trial were achievable, albeit at a more limited scale, while goal 3 was not.

An effective way to address the potential outcome of the trial from a modelling perspective is to conduct uncertainty analysis, something not previously attempted. Models are subject to equifinality, meaning it is possible to produce many equally well-calibrated models with different combinations of parameters. While the different combinations of parameters do not impact the estimates at the calibration observation sites, there can be significant differences in areas without observation data to constrain model parameters. Calibration constrained uncertainty analysis allows the effect of model equifinality to be assessed and presents a scientifically robust method to evaluate the possible range of effects from environmental stressors. This research used the results of the first two years (Phase 1 and

[1] A dry well is a well installed in the unsaturated zone to increase infiltration rates, while achieving attenuation of pathogens within the aquifer (Dillon et al. 2009).

[2] Where the outputs are fully determined by the parameter inputs and where an input will always produce the same output; contrasted with stochastic models that translate input uncertainty to output uncertainty.

2) of the pilot study to calibrate a base model, on which to conduct uncertainty analysis. The uncertainty analysis was then used to assess the MAR trials effect on the receiving environment.

1.1 Background

The study area totals approximately 66,000 ha of the mid-Canterbury region of New Zealand’s South Island. Specifically, the study area is bounded inland (northwest) by the foothills of the Southern Alps, by the Pacific Ocean to the southeast, with the Ashburton River/Hakatere to the northeast and the Hinds River/Hekeao to the southwest (Figure 1-1). The topography of the area is relatively flat and gently slopes to the sea with few distinguishable features, typical of the Canterbury Plains (Durney and Ritson 2014). The maximum elevation in the study area is approximately 350 m (Lyttelton vertical datum) near the base of the foothills, from where the land gently slopes down to the sea approximately 50 km to the southeast. In the coastward extent, the study area terminates in sea cliffs up to 10 m high.

The Valetta Irrigation scheme has heavily irrigated the area for nearly 70 years; however, in the past 20 years, farming practices in the study area have undergone significant changes. Dryland farms have adopted irrigation, and border dyke (flood) irrigated areas have converted to more efficient spray irrigation. Additionally, Ashburton District Council[3] has closed many of the stock water supply races that crossed the catchment, as users have found alternative sources of water.

Conversion to spray irrigation has led to the intensification of farming practices and a combination of increased demand for water with larger irrigated areas and a loss of incidental recharge (i.e. return flows) from inherently inefficient border dyke irrigation methods. It is estimated that as much as 2 m^3/s incidental recharge has been lost during the summer months as a result of the conversion to spray irrigation and the piping of the distribution network within the Valetta Irrigation scheme (Durney and Ritson 2014). Concurrent with these changes, foothills fed waterbodies across Canterbury have shown declines in flow; this has been attributed to a drying climate (pers. comm. Z. Etheridge 2017). All these factors have placed increasing pressure on the groundwaters between the Ashburton River/Hakatere and the Rangitata River. Groundwater levels have declined significantly since the early 2000s (Durney and Ritson 2014). Figure 1-2 shows the groundwater table decline in a typical Hinds well. As a result of the high allocation and usage of water, the Hinds Plains has been categorised as over-allocated regarding both groundwater quantity and quality (Durney and Ritson 2014; Scott 2013).

[3] The local territorial authority.

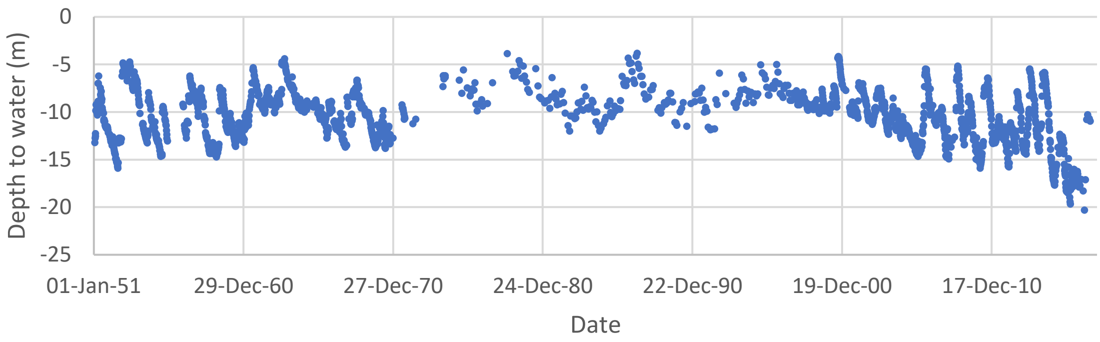

Figure 1-2 Depth to water in well K37/0215 Valetta.

Historically, the groundwater of the Hinds area supported an abundant lowland stream system that, until recent times, maintained vibrant biodiversity (Meredith 2006, 2014). In the last seven years, much of this lowland stream system has dried up, partly in response to three years of low rainfall (Golder 2017).

Groundwater concentrations of nitrate-nitrogen (NO_3-N) have risen simultaneously with the declining groundwater levels and spring-fed stream discharge (Scott 2013). NO_3-N levels have increased in response to the Hinds Plains becoming increasingly intensively farmed and as irrigation efficiency has improved (Dench 2017; Scott 2013). Since 2017, concentrations of NO_3-N in the lowland streams have exceeded 24 mg/l immediately following periods of intense rainfall (pers. comm. A. Meredith 2017). For perspective, the New Zealand national bottom line for aquatic ecosystem health has been set at a median concentration of 6.9 mg/l NO_3-N (New Zealand Government 2017).

The over-allocation of the Hinds groundwater led ECAN to instigate Plan Change 2 (PC2) of the Land and Water Regional Plan (LWRP) (Environment Canterbury 2017). As part of the PC2, MAR was put forward as one solution to address over-allocation of groundwater, regarding both quantity and quality limits (Golder 2017). As part of the planning process, all proposed plans must be reviewed and approved by a panel of independent hearing Commissioners who consider information put forward by the council and submitters on the proposed plan. The hearing Commissioners decided that evidence of the applicability of MAR in the Hinds catchment was needed and called for the adoption of a trial programme (Golder 2015).

Despite the Commissioners' call for a tangible demonstration of MAR before the adoption of a groundwater replenishment scheme, artificial recharge is not new in the Hinds catchment, albeit unintentional. The Valetta Irrigation scheme significantly raised groundwater levels during the period where the primary irrigation method was border dyke irrigation (Davey 2006a); similar findings were made by Dench (2017) in the area just south of the Hinds River/Hekeao. In part, this was one of the key reasons the community were so ready to adopt the concept of MAR (pers. comm. P. Lowe, Chair of the Hinds MAR Governance Group 2017).

Following the PC2 decision, ECAN, in partnership with the community and industry groups, began the design of the MAR trial. The primary site selection criteria were: the availability of land, depth to groundwater greater than four metres and availability of water supply. A review of potential recharge sites resulted in the selection of an unused land parcel at the corner of Frasers Road and Timaru Track Road at Lagmhor, owned by Ashburton District Council (Golder 2015). The MAR trial pilot study is presently 60% complete and thus far has shown promise in raising groundwater levels and improving groundwater quality.

1.2 Research Aim

The research aim is to estimate the impact of the Hinds MAR trial on groundwater levels, streamflow, water quality, and to assess whether the trial's stated goals are achievable. Analytical and numerical modelling is used to address this research goal.

Specifically, this research:

1. develops several analytical mounding models of the trial site, using the upper and lower bounds of estimated aquifer parameters;
2. develops a steady-state 3-dimensional numerical flow model of the trial site and surrounding catchment;
3. applies uncertainty analysis of the modelled parameters to assess the potential long-term effects of the MAR trial and, specifically, the estimation of 5^{th}, 50^{th} and 95^{th} percentiles and mean levels of groundwater mounding;
4. develops a transient transport model of the MAR trial.

This research is important because the Hinds MAR trial is acting as a test case for the national adoption of MAR as a remediation tool (pers. comm. B. Bower 2018). Even locally in the Hinds area, the trial is being used as a test case for the adoption of a catchment-wide groundwater replenishment scheme. Currently, there is no formally established method for assessing a potential site's worth or for estimating the effects of such a site. However, further trials are planned in the next few years. This research will help to inform the future approaches to MAR site assessment by demonstrating the use and effectiveness of uncertainty techniques in quantifying the probability of reaching target goals. Further, it will recommend the most practical modelling approach to the assessment of future MAR studies.

2 Literature review

2.1 Managed Aquifer Recharge

MAR is the intentional augmentation of an aquifer's recharge via controlled injection or infiltration of water into an aquifer system (Dillon et al. 2009). The following list is adapted from Metcalf and Eddy Inc. (2007) and Dillon et al. (2018) and while not exhaustive, details the primary purposes for MAR:

1. To reduce, stop or reverse groundwater declines.
2. To prevent saline intrusion into freshwater aquifers.
3. For harvesting and storage of surface water for later usage.
4. To prevent land subsidence.
5. To provide treatment of stormwater and or grey water to remove microbial contaminants for later recovery and utilisation.

Each purpose generally involves different implementations of MAR, with the two main technologies employed being injection of water via borehole and infiltration via infiltration/spreading basin (Ringleb et al. 2016). A third, less commonly used method is the dry well (described below). Other less widely employed methods, similar in approach to spreading basins, include enhanced stream bank infiltration and infiltration along dry reaches of losing rivers (Dillon et al. 2009; Wallbridge Gilbert Aztec [WGA] 2018). In Dillon et al. (2009), the various types of injection methods, such as aquifer storage and recovery from the same well, aquifer storage, and transfer and recovery from a different well are all considered under injection wells; while infiltration galleries, spreading basins, percolation tanks and recharge weirs are all considered as infiltration/spreading basin methods.

Injection wells involve the injection of recharge water directly into the saturated zone via the use of a purpose-designed borehole. Direct injection can be used in both unconfined and confined aquifer conditions (Metcalf and Eddy Inc. 2007). Key benefits of this technology are the relatively small footprint area of the injection wells. Injection wells can be used in areas where spreading basins cannot, for example, where there is insufficient land or in areas of highly variable topography. However, there are significant costs associated with this method: the construction of the borehole/well can be significantly more expensive than a spreading basin of similar capacity, the design life is shorter than that of spreading basins and maintenance more difficult and costly. Compounding the higher costs is the lack of soil aquifer treatment that is achieved with spreading basins (Kazner et al. 2012). The lack of soil aquifer treatment means that treatment of the recharge water is usually required before it can be injected into the aquifer. Other difficulties arise around the introduction of chemically different water into the aquifer, which can cause significant and unwanted geochemical reactions. The need for pretreatment and the possibility of geochemical reactions makes the use of injection wells significantly riskier than infiltration/spreading basins (Dillon et al. 2009; Metcalf and Eddy Inc. 2007).

Due to the cost and risks associated with injection wells, surface spreading basins are the oldest and most widely used methods of MAR (Metcalf and Eddy Inc. 2007). This method of artificial recharge involves the infiltration of water from the floor of the infiltration/spreading basin through the unsaturated zone. Infiltration basins are the preferred method of groundwater recharge because they are simple, make effective use of space and are easily maintained. Further, spreading basins provide the added benefit of soil aquifer treatment, as the water percolates through the unsaturated zone (Metcalf and Eddy Inc. 2007).

The third but less commonly used method of MAR is the so-called dry-well. The dry-well is a large diameter bore constructed in the unsaturated zone of the aquifer that allows injection of water to a specified depth, while still capturing the soil-aquifer treatment properties of the vadose (unsaturated) zone (WGA 2018). The design originated as a method of reducing the footprint area of infiltration/spreading basins to reduce land acquisition costs and to facilitate its adoption where insufficient land was available. The technique has also been used to penetrate deep enough into the unsaturated zone to bypass local impermeable layers (Metcalf and Eddy Inc. 2007; WGA 2018).

2.2 Assessing the effects of MAR

A review of the relevant literature suggests four main, not necessarily mutually exclusive, options are currently undertaken to evaluate the outcomes or effectiveness of MAR (Ringleb et al. 2016). These are:

1. detailed site-specific field investigations pre-implementation
2. laboratory studies for researching new methods
3. the adaptive or the “trial and error” approach
4. modelling.

Generally, only options 1 and 4 are pursued, with option 2 considered relevant for researching new recharge methods and option 3 limited to low-cost and low-risk rainwater harvesting projects (Ringleb et al. 2016). The Hinds MAR trial falls into options 1, 3 and 4. Initially, field investigations were conducted at two sites before a pilot-scale trial. These pre-implementation investigations indicated that significant recharge rates (0.5 m^3/s to 1.5 m^3/s) could be achieved (Golder 2015, 2017). After the trial commencement, these recharge rates were never achieved. The Hinds MAR trial is largely now of type 3, with monitoring data and analysis conducted after the fact, and the programme modified as problems were encountered.

The 2014 *California Department of Drinking Water guidelines for Groundwater replenishment using recycled water* also provide useful guidance as to the relative worth of assessment methods for MAR, as part of its assessment criteria (California Water Board 2014). The following relative worth multiplication factor rankings are taken from the legislation (California Water Board 2014, p. 19):

"*Method used to estimate the retention time to the nearest downgradient drinking water well (Virus Log Reduction Credit per Month*[4]*)*

- *Tracer study utilising an added tracer.* ***1.0 log***
- *Tracer study utilising an intrinsic tracer.* ***0.67 log***
- *Numerical modelling consisting of a calibrated finite element or finite difference models, using validated and verified computer codes used for simulating groundwater flow.* ***0.50 log***
- *Analytical modelling using existing academically accepted equations, such as Darcy's Law to estimate groundwater flow conditions based on simplifying aquifer assumptions.* ***0.25 log***"

Each of the respective assessment methods are used to estimate the retention time of the recharge water before it reaches a point of interest. Treatment of the recharge water must be designed to meet a required log reduction in pathogen count, such that the observed pathogen count is below its set limit at a specified distance from the recharge site. The multiplication factor based on the assessment method means that if 3 log reductions in pathogen count are required and the assessment (via numerical modelling) showed three months residence time, there would be a requirement for an additional 1.5 log reductions via the treatment of the recharge water. If on the other hand, a tracer study had been used and showed a three-month retention time, no additional treatment would be required.

This ranking of the relative worth of various analysis methods shows that the weight of evidence is given to field observations. Numerical modelling comes third as an assessment method but ranked ahead of analytical modelling. While numerical modelling can never replace observations, it is still useful as a method to extrapolate observations either spatially or temporally.

Despite the limitations and uncertainty inherent in numerical modelling, it is becoming increasingly necessary, because as a predictive tool, it enables extrapolation of limited monitoring data. Additionally, it is significantly cheaper than large scale high-intensity monitoring or intensive in-situ pre-testing as a means of assessing the more extensive effects of MAR programmes. When modelling is undertaken, it is important to provide decision-makers with the most useful predictive information (Ringleb et al. 2016).

[4] The credit system is designed for the MAR planning stage and works as a reduction factor on the residence time (in months) of water entering the groundwater system via a MAR. It guides the planned level of water treatment required in terms log reduction of biological contaminant count based on the pre-construction assessment method.

The modelling techniques used in understanding or quantifying the effects of MAR make use of methods used in the rest of the hydrological industry. Historically, those evaluating the impact of MAR programmes have relied on deterministic approaches that can risk presenting a biased outcome regarding the proposed aquifer replenishment programme (Maliva 2015).

In recent times, the hydrological industry has been moving towards the adoption of advanced uncertainty analysis and the quantification of uncertainty. It is important that modelling of MAR outcomes, likewise, adopts uncertainty analysis to give decision-makers around the adoption of MAR, a balanced perspective with regards to its likely success (Maliva 2015). Full uncertainty analysis involves addressing all sources of uncertainty, including informational (uncertainty associated with observation data), structural/conceptual and parameter uncertainty. To date, no means exist to investigate all of a models structural and conceptual uncertainty (short of building and calibrating multiple different model structures), with the techniques generally limited to parameter uncertainty. Despite the limitations of the present uncertainty techniques, the probabilistic model results are still less biased than those from deterministic models alone (Maliva 2015).

2.3 International use of modelling to assess MAR

A review of modelling applications related to MAR by Ringleb et al. (2016) summarised the approaches to modelling MAR in 216 international case studies. The most commonly adopted modelling platform is MODFLOW (71 instances).

Thirteen key categories of MAR modelling studies are identified (Ringleb et al. 2016):

1. **Feasibility –** to determine whether a site is suitable.
2. **Design** – optimal design of the recharge mechanism.
3. **Optimisation** – assessing the optimal schedule/timing of recharge and recovery operations.
4. **Recovery efficiency** – quantification of the recovery efficiency for aquifer storage and recovery (ASR) programmes.
5. **Residence time** – the residence time of the recharged water, e.g. to ensure adequate soil aquifer treatment.
6. **Geochemical processes** – to understand any geochemical processes that may occur due to the infiltrated/injected water, e.g. sorption/adsorption.
7. **Clogging –** to assess the rate at which the recharge system may become clogged.
8. **Water quality** – investigations into possible water quality changes resulting from the recharge programme.
9. **Soil aquifer treatment** – assessment of the retardation/attenuation of organic/microbial organisms in the aquifer.

10. **Groundwater management** – assessment of water level and volume responses to the recharge programme.
11. **River flows** – assessment of any changes to river flows because of the recharge programme.
12. **Risk assessment** – assessment of potential hazards associated with the recharge programme, e.g. groundwater flooding and pathogen transmission.
13. **Saltwater intrusion** – the assessment of the effectiveness of MAR in preventing saltwater intrusion.

While recharge basin methods are the most common form of MAR, most modelling studies have focused on injection wells and ASR. Where studies have concentrated on infiltration basins, most have been focused on groundwater management, with a secondary focus on water quality changes (Ringleb et al. 2016). Most modelling falls into the planning stage of the design and construction of recharge programmes with secondary consideration given to predicting the long-term effects of MAR on the receiving environment (Ringleb et al. 2016).

The present research is in the groundwater management and water quality category, focusing on groundwater level and water quality changes expected by the completion of the trial. Emphasis is placed on quantifying the range of probable outcomes using uncertainty assessments, which Maliva (2015) identifies as a significant gap in previous MAR modelling studies.

2.4 Previous MAR studies in Canterbury, New Zealand

MAR is mainly seen as a solution in semi-arid environments (Rahman et al. 2012), and as such, artificial recharge is not new in Canterbury. During the 1990s, three trials were carried out near Christchurch City (Moore 1994; Sinclair Knight Merz 2009) and another near the town of Rangiora using the Eyre River (Pattle Delamore & Partners 2007). Like the present Hinds MAR trial, the Christchurch-West Melton experiments used water delivered by water supply races that traverse the Canterbury Plains. In these instances, however, water was discharged into soakage pits rather than a larger infiltration basin. The experiments conducted during the 1990s highlighted clogging would be a problem long-term due to siltation (Moore 1994). Only one of the three trials near Christchurch City and the Eyre River study showed groundwater rise across a wide area, with the others resulting in only limited responses. In all cases, despite small recharge area footprints, recharge rates exceeded 0.1 m^3/s (Sinclair Knight Merz 2009). The Eyre River study was the most successful of these studies using 2.7 m^3/s water from the Waimakariri Irrigation Ltd scheme and discharging it into the Eyre River for three weeks. Groundwater levels rose significantly (over 2.5 m) across more than 40 km^2 and remained elevated for two to three months after the trial (Pattle Delamore & Partners 2007).

2.5 The Hinds MAR trial

The Hinds MAR trial infiltration basin consists of an approximately 1 ha recharge basin and forebay at the intersection of Timaru Track and Frasers Road. Water for the trial is sourced from an Ashburton District Council stock water supply take on the Rangitata River and delivered to an on-farm storage pond at the end of the Valetta Irrigation Scheme via a piped network, from where it travels via an open race to the recharge basin (Figure 2-1).

The infiltration basin forebay is intended to act as a settling pond to remove suspended sediment and thereby reduce the longer-term clogging risk at the site (Figure 2-2) (Golder 2017). Following a review of recharge options and study of the receiving environment, which indicated approximately 20-30 m of unsaturated zone material beneath the pilot site, a combination of an infiltration basin and 28 approximately one-metre wide and six-metre deep unlined dry wells were selected for construction. The dry wells were excavated by clamshell and backfilled with clean boulders to increase the possible infiltration rates (Golder 2017).

The construction of the recharge basin involved scraping of topsoil at the lowest elevation point and incision into the landscape at the highest elevation point. Excavated material was stored onsite as a large bund at the north-eastern end of the basin.

Figure 2-1 Distribution race between Valetta Scheme Pond No. 3 and the MAR recharge basin (adapted from Golder 2017).

Figure 2-2 Hinds MAR infiltration basin (from Golder 2017).

A groundwater monitoring array was developed to track both the mounding and clean water plume to determine the efficiency of the recharge trial. Instantaneous flow measurements were recorded at two locations on the distribution race that supplies the recharge basin, to enable determination of conveyance losses. Groundwater level data was collected at the 22 wells shown in Figure 2-3. Seven of these wells were purpose drilled (alpha-numeric classification starting with BY) for the trial while the rest were existing sites.

During the second year of monitoring, the number of observation-locations was reduced to the seven purpose drilled wells and ECAN monitoring well K37/1748[5]. Electronic data loggers, recording water pressure and temperature at 15-minute intervals were used in 20 of the 22 sites. Atmospheric barometric pressure was also recorded at the same frequency at the trial site. Water quality sampling was carried out at all the monitoring wells on at least one occasion, with most sampled at least monthly. Samples were analysed for a suite of biological and chemical properties including dissolved oxygen, NO_3-N, dissolved reactive phosphorus (DRP) and Escherichia coli bacteria (E. coli). In addition to the groundwater monitoring, five surface water flow recorders were installed in the local drains to detect any changes in flow resulting from the trial and three existing sites were utilised (Figure 2-3) (Golder 2017).

---

[5] Environment Canterbury uses an alpha-numerical system for identifying wells within its jurisdictional boundaries. The initial letter and number refers to the identification code of New Zealand topographic map covering the region in which the well is located, while the final number sequence refers to the order in which the wells were registered with the council.

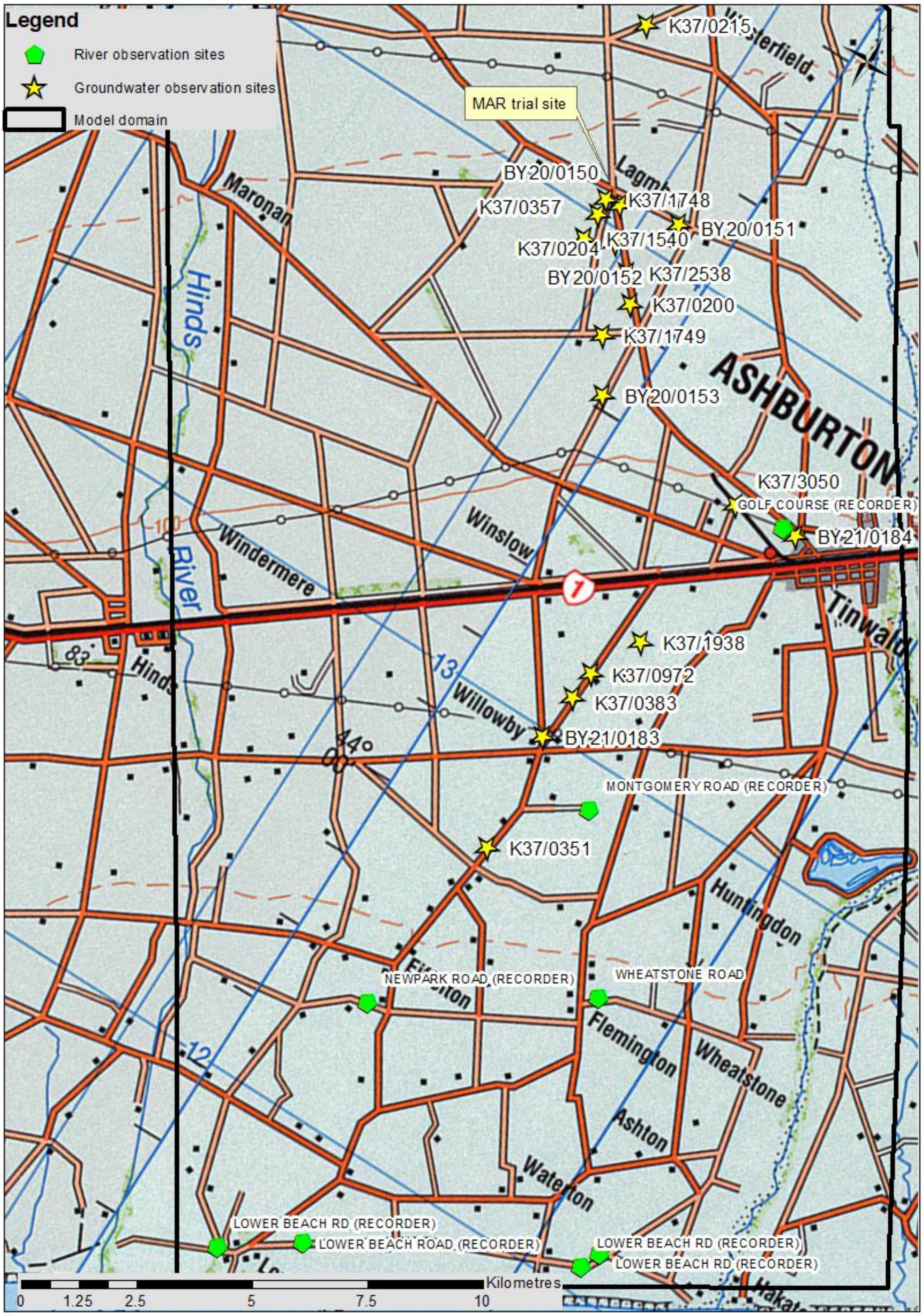

Figure 2-3 Hinds MAR trial year one monitoring array.

The first two phases of the Hinds MAR trial ended in June 2018. The first year of operation was intended to demonstrate that water could infiltrate into the ground in significant enough volume to cause a measurable change in groundwater level and quality (Golder 2017). Phase 1 showed a definite rise in groundwater levels and corresponding decreases in measured groundwater NO_3-N levels for several kilometres down-gradient (Golder 2017). The second years results demonstrated that the clean water plume from the trial site continued to flow towards the coast, generally following the topography of the region (WGA 2018). Figure 2-4 illustrates the propagation of the clean water plume by the end of year two, based on field measurements and interpolation.

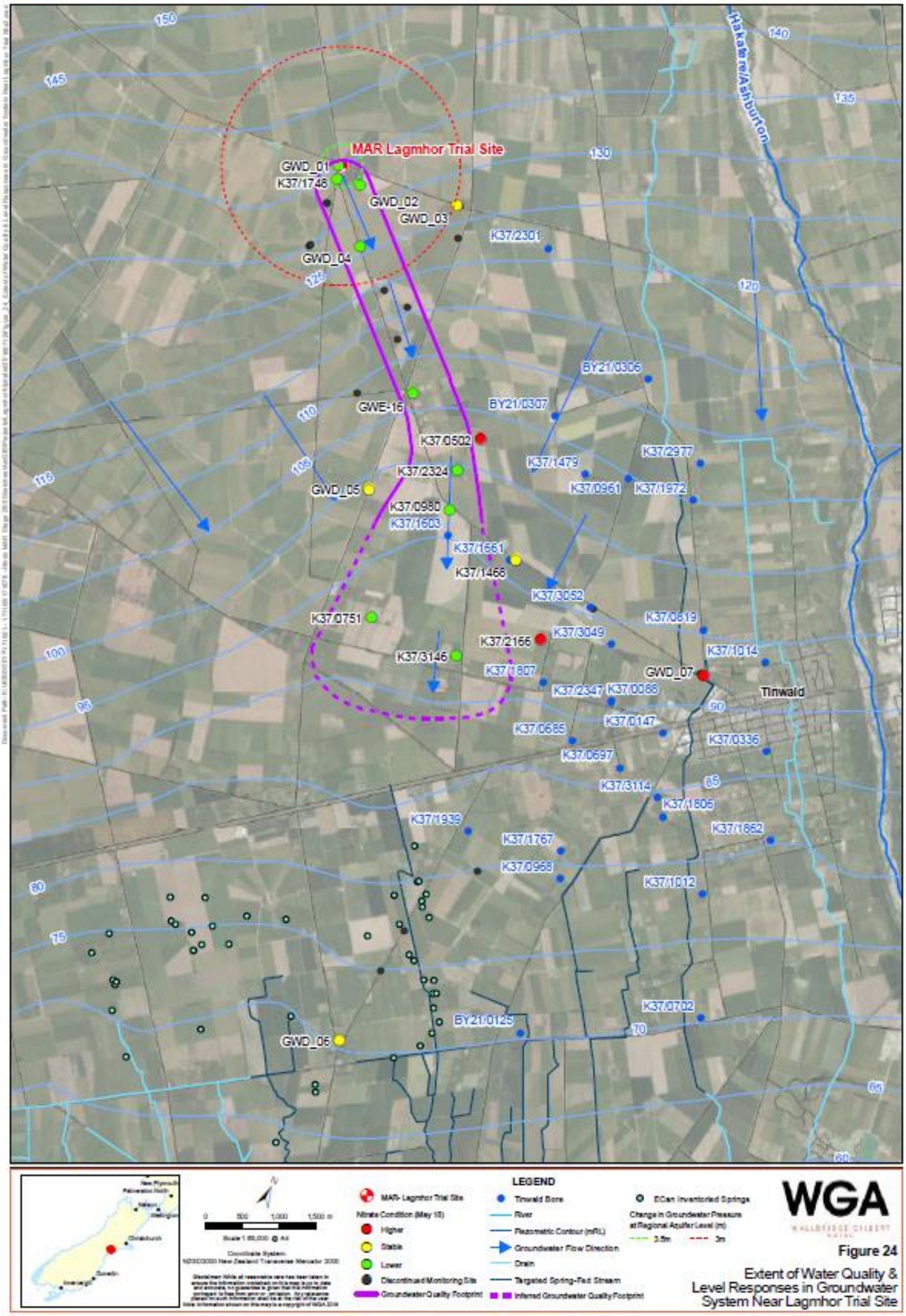

Figure 2-4 Clean water plume extent (purple lines) (from WGA 2018).

The second year of the trial included the installation of an additional clamshell excavated dry well inside the recharge basin to induce higher recharge rates. This hole was designed as a dry well and drilled to 18 metres depth. While the overall year two recharge rates were higher than those from year one, reaching up to 0.14 m^3/s, the findings suggest that the additional dry well contributed little to the improvements. During the second year of the trial, WGA had also conducted further tests in the recharge basin to optimise recharge rates. The tests involved assessing the effects of various water level stage heights (and by proxy; ponded water depth) in the infiltration basin. Testing involved adjusting the depth of the water in the recharge basin and recording inflow rates. WGA attributed the second-year improvements in recharge rates to the optimised stage height rather than the additional dry well (pers. comm. B. Bower 2019).

2.6 Previous modelling of the Hinds MAR trial

Three modelling investigations have been conducted previously to investigate MAR in the Hinds catchment. Durney et al. (2014) conducted the first 3-D transient numerical modelling (using MIKE SHE software) to investigate the feasibility of MAR, as a regional solution to declining water levels in the area. The modelling suggested up to 7.5 m^3/s artificial recharge would be required to maintain and restore groundwater levels. Furthermore, compensation would be necessary for the loss of incidental aquifer recharge that would result from piping the Mayfield Hinds Irrigation Scheme distribution races.

Following the decision to conduct the MAR trial at Lagmhor, Golder conducted two infiltration tests to estimate recharge rates at the trial site. Golder (2015) modelled the test using SEEP/W software and estimated between 0.5 m^3/s and 1.5 m^3/s possible infiltration across the 9500 m^2 recharge basin. The Hinds MAR trial demonstrated that these predicted recharge rates were far higher than what was achievable (~0.1 m^3/s). Kronast (2016) remodelled the infiltration trial in SEEP/W, building a model to the scale of the actual trial recharge basin. Previous estimates had been based on upscaling the recharge rate estimates from much smaller soakage pits, rather than modelling the infiltration basin footprint with the parameters derived from the soakage pit modelling. Kronast (2016) estimated a significantly lower infiltration potential of between 0.115 m^3/s and 0.460 m^3/s and theorised that the initial infiltration potential estimates were overly high because of upscaling of the infiltration test recharge rates. Kronast (2016) recommended that instead, the recharge potential should have been estimated by modelling the actual basin size using the infiltration test parameters.

Golder (2017) presented a deterministic calibrated model using a 10-layer transient MODFLOW-NWT and MT3DMS model, that extended approximately 10 km inland and coastwards of the recharge basin. The model grid resolution was 250 m by 250 m, with each layer 15 m thick. This model was used to extrapolate the first-year (first 6 to 9 months of data) results to the end of the trial. They found that the plume of clean water would likely reach State Highway One (SH1) approximately 7 km down-gradient by the trial's completion (Figure 2-5). However, their work was hampered by the location of the model's southern boundary, which was too close to the trial site and impacted both the mounding and water quality results (Golder 2017).

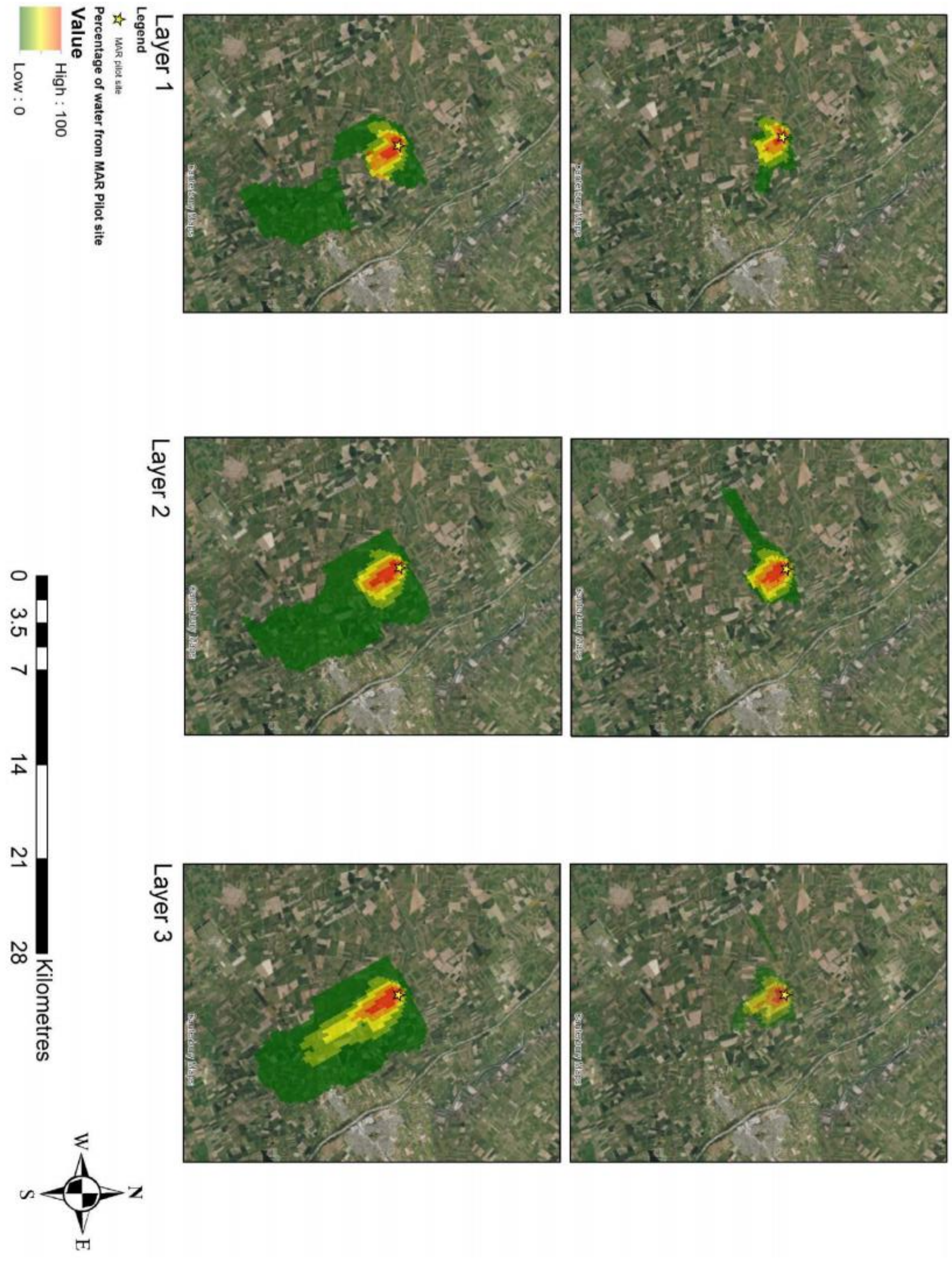

Figure 2-5 Results of MT3DMS water quality modelling (from Golder 2017). The top row of images shows the results after one year, while the second row shows the predicted results after five years.

2.7 Uncertainty

In the context of groundwater modelling, uncertainty refers to the variability in model results caused by the imperfect representation of the processes in the model and errors in available observation data. Put another way; uncertainty arises both because of errors in observations and the inherent inability to capture all the physical processes that are occurring. All models are simplifications of the systems they are trying to represent, and due to the imperfect mathematical representations of the real world within models, no single model exists that can perfectly describe the natural environment (Beven and Binley 1992). Additionally, there is natural variability in observational data; where without having the actual population of observation data that captures the real statistics of the data set, all inferences drawn from the data are uncertain as they will not represent the actual range of values likely to be encountered (Cozby and Bates 2012; Loucks and van Beek 2005).

Loucks and van Beek (2005) propose a classification based on the source of uncertainty. In this system of classification, key categories are:

1. informational uncertainty, which is the lack of knowledge about real boundary conditions, lack of correct initial conditions, error in measurements and lack of understanding of the observation's true statistical population.
2. model uncertainty, which is uncertainty introduced from the choice of model structure and boundary conditions, lack of knowledge of actual physical parameters (parameter uncertainty), and variability at a scale smaller than the resolution of the model.

Quantifying the impacts of informational uncertainty (beyond measurement error) is difficult without an easy mechanism to build and test multiple models. The same difficulties are encountered when assessing the effects of model structure uncertainty. Presently, no easily adoptable approach allows all sources of structural uncertainty to be addressed. The assessment of uncertainty in this research is limited to model uncertainty associated with model parameters.

PEST (short for Parameter ESTimation) (Doherty 2003) is used here to assess the predictive model uncertainty. PEST enables several approaches to uncertainty analysis. Two PEST based methods, (linear analysis and calibration constrained Monte Carlo) are discussed in the following section, along with Generalised Likelihood Uncertainty Estimation (GLUE). The other PEST uncertainty methods not discussed are predictive calibration and predictive maximisation/minimisation (Doherty 2003).

2.7.1 Linear analysis

The basis of linear uncertainty analysis is that the model being used is a matrix that acts on a series of parameters, to generate modelled outputs using equations that maintain linearity in the model calculations. Linear analysis requires that the prior parameter probabilities and measurement errors follow Gaussian distributions. The matrix assumption requires that some parameters are associated with observations via model calibration, while others are estimates required for the model to function (Doherty et al. 2010).

A benefit of linear analysis is the ability to investigate the significance of any single parameter or observation regarding the uncertainty it introduces into the model. The analysis is undertaken using two hypothetical cases: one assuming that the parameter/observation being investigated is an unknown variable and the other that it is a fully known variable. By comparing the difference in model results, under both cases, it is possible to determine the significance of the parameter/observation. Further, it enables things not commonly seen as a parameter to be investigated, such as the impact of model boundaries, thereby allowing some level of structural uncertainty to be assessed (Doherty et al. 2010). An example of this use is the investigation of data worth, e.g. if observations are available in a specific area, how much of an impact does that have on the uncertainty associated with the model predictions. Additionally, it can be used to highlight the relative worth of each parameter in terms of reducing uncertainty (Doherty et al. 2010). However, compared to other methods of assessing uncertainty, the approach is less useful in estimating the alternative outcomes the parameter uncertainty causes.

Of the three approaches to uncertainty analysis discussed here, linear analysis has the most constraining assumptions. While the assumptions introduce more limitations than the other methods, the technique is still beneficial as it can provide approximations of the actual parameter and predictive uncertainty (Doherty et al. 2010). Further, increasing the usefulness of the method is its independence from the model parameter values and outcomes.

2.7.2 Objective function constrained Monte Carlo simulation analysis

Objective function constrained Monte Carlo (Monte Carlo approach) uncertainty analysis is perhaps the easiest approach to understand when it comes to uncertainty analysis.

The "objective function" is a lumped calibration statistic used to determine model performance against observations. Unlike an optimisation model where the objective function is minimised to reach a point where the model is considered calibrated, the approach involves generating large numbers of random parameter sets and then running simulations of each set (Loucks and van Beek 2005).

The real utility of the Monte Carlo approach is where the parameters being generated are random inputs that will produce random model outputs. Random inputs are any observation or parameter value that cannot be predicted with any real certainty; in effect, these can include many environmental observations/parameters, such as rainfall, river stage, groundwater level, hydraulic conductivity, and streambed conductance (Loucks and van Beek 2005).

The model is then run using the randomly generated parameter sets, with the outputs of each parameter set assessed against observations to determine the relative goodness of fit using an approach like GLUE

(Stedinger et al. 2008). GLUE is a method of estimating uncertainty that allows the rejection of any model that has results significantly different from observations. While the GLUE approach may reject some sets as being too divergent from the observations, ultimately the modeller ends up with many equally plausible models from which parameter and predictive uncertainty may be assessed. The ensemble of models produced may then be used to assess the probability of a specific scenario coming to pass. While this approach can present a robust test of uncertainty, the computational burden of generating enough simulations to produce meaningful results makes its applicability limited to simpler models (Loucks and van Beek 2005).

2.7.3 Null Space Monte Carlo

Null Space Monte Carlo (NSMC) was designed to get around the computational burden associated with Monte Carlo analysis while providing a more realistic approximation of predictive uncertainty than linear analysis.

NSMC is a unique approach adopted in PEST and relies on a model having been previously calibrated using the Singular Value Decomposition (SVD) assisted procedure. SVD assisted calibration is an approach that enables faster calibration of highly parameterised models than would generally be possible with PEST. It works by reducing the number of parameters to a smaller group of super parameters through SVD. SVD does this by producing smaller linearised approximations of the models underlying model matrices. The objective of SVD is to determine which parameters can be estimated in the model and which cannot. SVD identifies combinations of observations that provide unique information about specific combinations of parameters. In doing so, the approach also identifies both those parameters to which observations are insensitive and those combinations of parameters that when varied in conjunction with each other, offset their effects on the model (Doherty and Hunt 2010). Those parameters or sets of parameters that are uniquely estimable are known as the solution space, while those parameters or combinations of parameters that are not uniquely identifiable are called the “null space”.

While every parameter in the “null space” may have no unique effect on the objective function or observations to which the model was calibrated, it may affect what occurs where there are no observations (Doherty et al. 2010). NSMC works by taking the SVD assisted calibrated model and changing the “null space” parameters, allowing the modeller to produce many stochastically generated calibrated models. While the solution space parameters often need adjusting to bring the new models back into calibration, as these are much smaller in number than the overall parameter set, the recalibration is relatively quick (Doherty et al. 2010).

2.8 Synthesis

This review has identified the trend in international research to utilise uncertainty analysis in assessing the potential effects of environmental stressors, so that decision-makers are provided with unbiased assessments. This need carries over to the assessment of MAR programmes (Maliva 2015). To date, the Hinds MAR trial has only been assessed using deterministic models that have not explored the full range of possible outcomes. Suitable uncertainty analysis techniques have been identified that are readily available and standard practice in other areas of water resource studies. This research used uncertainty analysis tools to aid in understanding the probable outcome of the Hinds MAR trial.

3 Hydrogeological conceptualisation

3.1 Conceptual model

A conceptual model describes the key components and understanding of the hydrological processes affecting the resource to be modelled (Enemark et al. 2018). Viewed at the regional scale, this study area's conceptual model is relatively simple, consisting of a large unconfined gravitational flow system, crossed by several hill fed streams that lose water in their upper reaches before gaining flow in the lower reaches.

Contrary to the conceptual simplicity, in detail, the aquifer system is highly complicated, being both horizontally and vertically stratified. This means the aquifer parameters vary both vertically and horizontally. The vertical stratification leads to high vertical anisotropy and semi-confined conditions with various levels of confinement when analysed in traditional analytical models. However, as the aquifer material was deposited by various high-energy river channels that are discontinuous and randomly truncated and not in a low-energy environment, there are no true continuous horizontal layers or vertically confining materials to lead to such a classification. Instead, the coarser river channels are interfingered with associated outwash gravels and finer sediments that would have been deposited along the channel margins, both horizontally and vertically. When the aquifer is considered in bulk, the depositional regime has produced very high vertical anisotropy and occasionally high horizontal anisotropy.

3.2 Climate

The climate in the study area is similar to the rest of the Canterbury Plains. Storm fronts from the south to easterly directions provide the dominant source of rainfall. Rainfall generally ranges from approximately 600 mm/yr at the coast, to 950 mm/yr at the base of the foothills (Durney et al. 2014). Most months receive roughly the same amount of rainfall when a long enough period is considered, e.g. 40 years (Figure 3-1). No measurements of potential evapotranspiration (PET) are taken within the study area, though three nearby sites maintained by the Meteorological Service of New Zealand Ltd (MetService) provide a continuous record. The MetService sites show that PET has a markedly seasonal pattern, with the highest levels recorded during the summer months, when it usually exceeds rainfall (Durney et al. 2014). Figure 3-1 shows that there is significant seasonality to PET, unlike rainfall. However, the variation in any given month is significantly less (Durney et al. 2014).

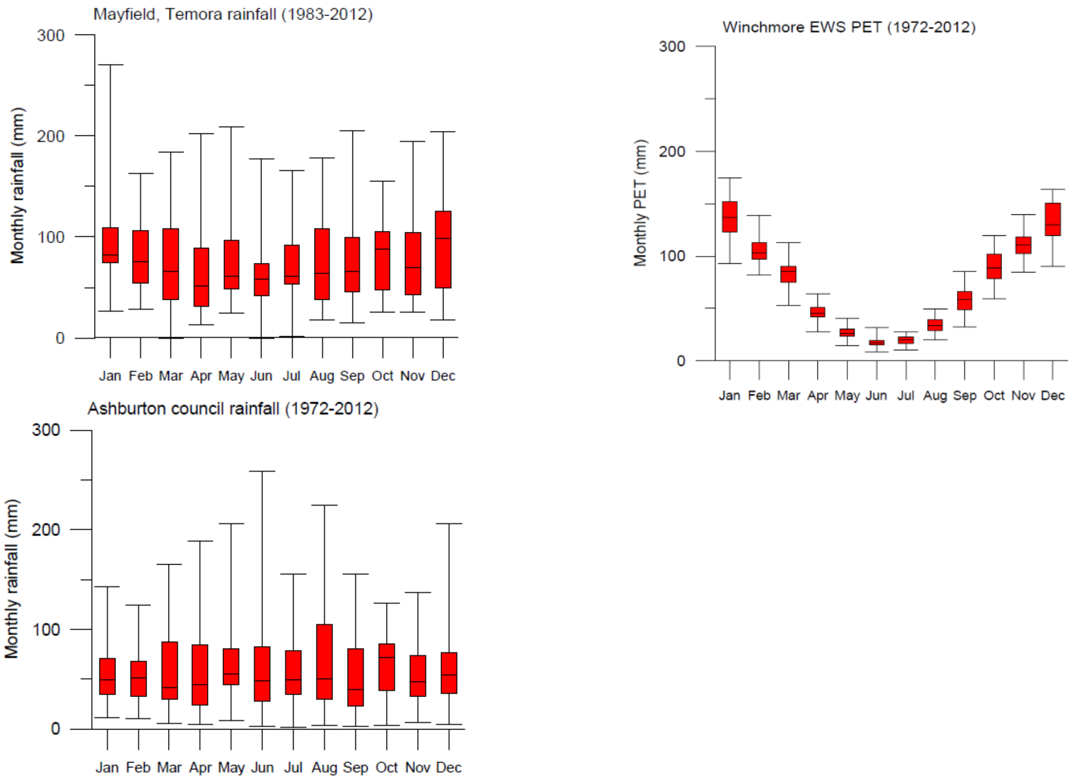

Figure 3-1 Box plot of the variation in monthly and seasonal rainfall and Penman evapotranspiration within the study area (from Durney et al. 2014).

3.3 Rivers

The Ashburton River/Hakatere and Hinds River/Hekeao are the main rivers in the study area and form the north-eastern and south-western boundaries, respectively.

The Ashburton River/Hakatere consists of two main branches, the North Branch and the South Branch. Both branches have their headwaters on the eastern side of the Southern Alps divide. The South Branch is glacial fed, with strong perennial flows, while the North Branch is a smaller foothill fed river that dries in its middle reaches. The total catchment area for both branches is approximately 1372 km^2, with the South Branch making up the dominant component of this (~909 km^2). The two branches extend across approximately half of the Ashburton plains before joining just inland of SH1 (Boyle 2012).

The South Branch acts as the natural boundary for the study area. Gaugings conducted by ECAN and its predecessor, the South Canterbury Catchment Board, show that between the flow recorder station at Mt Somers and the Valetta Bridge (Figure 3-2), the river loses flow to groundwater. This section of the South Branch shows that losses are flow-dependent and range from approximately 0-1.5 m^3/s, with the latest analysis, suggesting that on average, the South Branch contributes about 0.5 m^3/s (Durney 2019). Below the Valetta Bridge to the confluence with the North Branch, the river regains much, if not all the flow lost (Durney and Ritson 2014), and on average approximately 0-1.5 m^3/s is regained (likely from Bowyers and Taylor's streams).

Downstream from the confluence, gauging data suggests little in the way of gains or losses in the flow. It is not until the last couple of kilometres before the river discharges to the coast that anything is discernible in the way of interaction with groundwater (Durney et al. 2014). In the final two kilometres before the sea, the river once again appears to lose flow to groundwater, based on ECANs concurrent gauging data.

Additional to the main rivers flow, water quality was measured at five dedicated surface water sites and three existing surface water sites during year one of the trial.

Figure 3-2 Upper catchment recorder sites on the Ashburton River/Hakatere.

The Hinds River/Hekeao is a minor foothill fed river that initially consists of two branches that join at the top of the Ashburton plains. The total catchment area is approximately 350 km^2 (Durney et al. 2014). The river is intermittently flowing and loses all of its flow to groundwater shortly after its confluence. During the summer months, most of the length of the river is dry. Much farther downstream, around SH1, the river gains significant inflows from groundwater. From just coastward of SH1, the river flows perennially (Durney and Ritson 2014). Median inflows from the north and south branches of the Hinds River/Hekeao are 0.16 m^3/s and 0.25 m^3/s for the south and north branches respectively. Median flow at Poplar Rd below SH1 is 0.5 m^3/s.

3.3.1 Coastal Streams

Many spring-fed streams occur northeast of SH1 and constitute a significant portion of the groundwater discharge in the zone (other discharges include pumping and offshore flow). Flows in the coastal streams have largely dried up in the last two decades (Durney and Ritson 2014; Golder 2017). These streams, in conjunction with the Hinds River/Hekeao, form the remnant of the historic swamp (Mitchell 1980). Historically the spring-fed streams (including those south of the Hinds River/Hekeao) discharged as much as 7 m^3/s from groundwater. Limited gauging and recorder data combined with field inspection suggest that across the 2015 to 2017 period the spring-fed streams in the study area may have still discharged between 1 m^3/s and 2 m^3/s, though due to data sparsity, this is more a qualitative than quantitative estimate. One of the main aims of MAR is to help restore flows to the spring-fed streams.

3.3.2 Stock water supply and distribution races

Inside the study area, the Ashburton District Council stock water supply network takes water from several sources, but primarily from the Brothers Intake on the Ashburton South Branch. Water usage data from 2012 to 2017 show the Ashburton District Council took between 700 l/s and 1500 l/s on average via the Brothers Intake. The length of the active supply races represented in the model is approximately 260 km; over this length, the races lose between 80% and 90% of their flow to groundwater (Hall 2012).

3.4 Geology

Burnham Formation alluvial gravels laid down in the last glaciations[6] dominate the study area. The Rangitata River in the south and the Ashburton River/Hakatere in the north are the primary sources of

[6] Otiran Glaciations that ended approximately 14,000 years before present.

these sedimentary deposits. This geology consists of both massive and poorly sorted gravels with a silty sand matrix (Barrell et al. 1996).

The study area is made up of two river fans; the Rangitata in the south and the Ashburton in the north (Figure 3-3). The two river fans are expected to interfinger at depth. Most of the sediment was derived from the Rangitata River, as indicated by the topography of the Hinds Plains that slopes away from its highest point near the Rangitata Gorge. During the glacial periods, flows were restricted by glacial lakes and ice, which diverted some flow from the South Ashburton River/Hakatere through the north branch of the Hinds River/Hekeao. Periodic glacial lake breaches are believed to have occurred, sending large volumes of water and sediment offshore (Barrell et al. 1996).

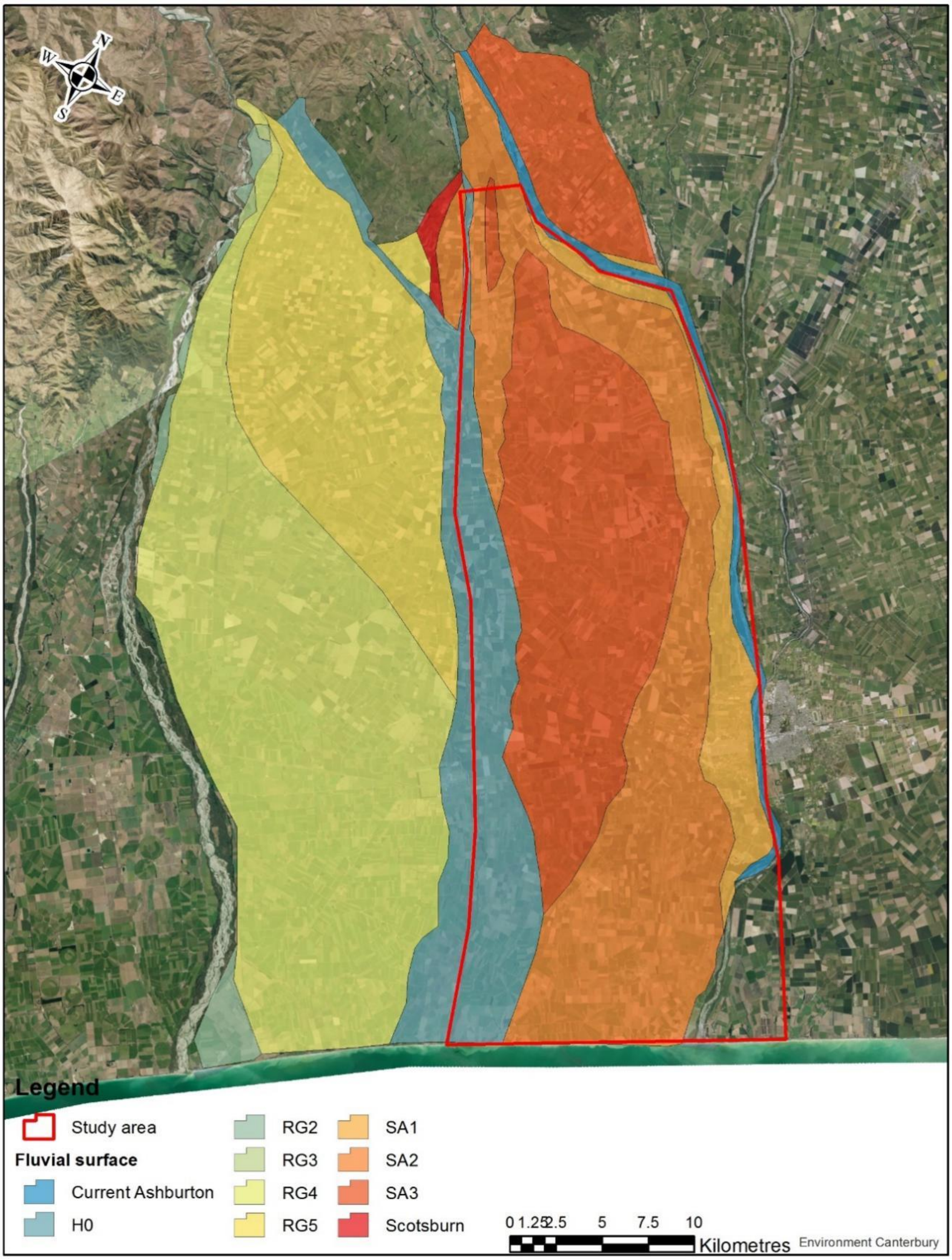

Figure 3-3 Fluvial surfaces from the Otira glaciation. RG refers to the Rangitata River fans, while SA refers to the South Ashburton River/Hakatere fan deposits (adapted from Barrell et al. 1996).

Barrell et al. (1996) describe how the sedimentary formations sit unconformably on Palaeozoic to Mesozoic age greywacke that forms a substantially impermeable flow boundary. The sedimentary formations have been mapped using gravity surveys as much as 1 km thick at the Rangitata Huts, with Quaternary sediments alone more than 400 m thick. Analysis of bore logs suggests that at the Ashburton River/Hakatere mouth, the previous interglacial deposits are encountered approximately 30 m below surface and contain higher fines content, perhaps acting as an aquitard (Brown et al. 1988).

In the coastal region, there are peaty and silty loams up to approximately one metre thick formed during the Holocene; these are interfluve swamp deposits laid down by the Hinds and Ashburton Rivers (Durney et al. 2014; Mitchell 1980).

Offshore the sedimentary deposits extend approximately 80 km from the coast and are up to 1.5 km thick. Offshore deposits are primarily of sand, silts and marine muds containing some relic gravel deposits (Barrell et al. 1996). Presently, no known gravel deposition is occurring offshore with the primary deposits being fine and very fine sands. The gravels carried offshore by the rivers and eroded from the coast are carried north by the longshore drift along the Canterbury Bight (Barrell et al. 1996).

Grey gravel layers/channels up to two metres thick are present in the youngest Q1a formations (Figure 3-4) (current fluvial deposits); these gravels are unweathered and lacking in matrix. Along the coastal cliffs, matrix dominated grey-brown sandy gravels are exposed (Q2a [SA2 and RG4]). The Q2a formation is more weathered than the Q1a and often has iron staining; the matrix also tends to show signs of weathering. Apparent in several locations are half metre-thick peaty loess bands that span several hundred metres. In the upper parts of the plains, the brown sandy gravels may be as much as 50 m thick, with layers of various clast sizes present. The grey-brown gravels represent the Otira glacial deposits and thin toward the coast where their minimum thickness is approximately 10 m. Beneath the grey-brown gravels lie brown gravels readily identifiable in the terraces opposite the inland Rakaia Bridge. These older brown gravels have been correlated with the Woodlands Formation (Barrell et al. 1996).

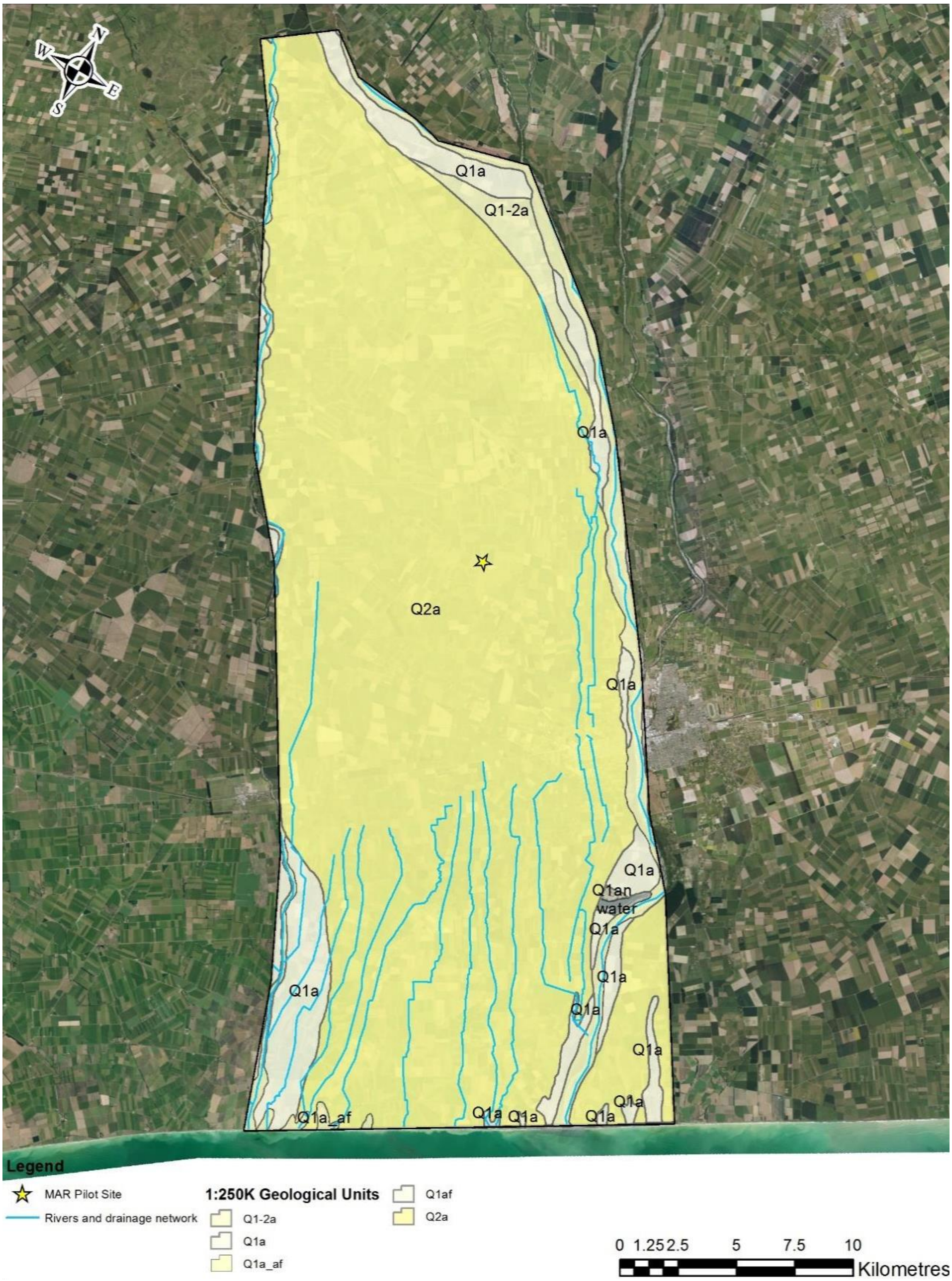

Figure 3-4 GNS Q-MAP geology of the study area.

3.5 Hydrogeology

As with much of the Canterbury Plains, most groundwater flow appears to occur as preferential flow in a small fraction of the aquifer material. Research conducted on the unconfined section of the Canterbury plains aquifer near Burnham, Canterbury, found that more than 90% of flow occurs in the preferential flow channels (Dann et al. 2008). After conducting fieldwork in the area investigating and mapping quarries and coast outcrops, Davey (2006b) characterised the geology of the Hinds aquifer system in much the same way as Dann et al. (2008). Figure 3-5 presents the idealised block diagram of the Canterbury (including Hinds) Plains (Davey 2006b).

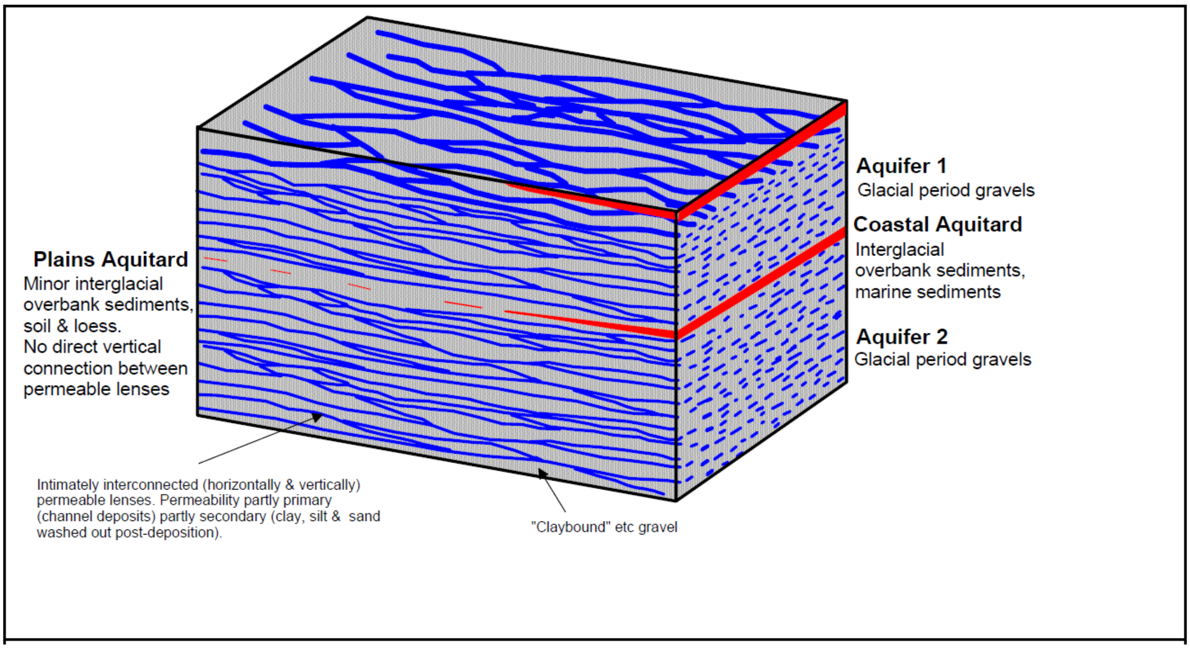

Figure 3-5 Davey's (2006b) conceptualisation of the Canterbury Plains aquifer system.

Durney and Ritson (2014) conceptualised the hydrogeology of the study area as a gravitational flow system. The less permeable strata within the aquifer act as an aquitard and cause a transition from unconfined to increasingly semi-confined (leaky) conditions. Groundwater isotope data collected by ECAN shows that recharge infiltrates rapidly to some depth (Hanson and Abraham 2013). Rapid infiltration in aquifers with strong downward hydraulic gradients, such as the Canterbury Plains, is expected.

Based on the previously described geology and hydrological responses, it is proposed that the aquifers of the Hinds Plains are a singular stratified unconfined aquifer. Essentially this means that the hydrological properties, such as hydraulic conductivity, specific storage and specific yield vary both horizontally and vertically because of the depositional sequence and environment. The less permeable strata within the aquifer act over the short-term as an aquitard, leading to the misconception that the aquifer is semi-confined (leaky) when aquifer tests are analytically analysed. The stratified nature introduces complications when assessing the effects of recharge and pumping, as any parameters that are estimated are an equivalent bulk value relative to where they are estimated. Preferential flow at the local scale suggests the aquifer is essentially dual-domain with most of the flow occurring as conduit flow. Generally, the simplest approach to assessing the aquifer properties in such a geological setting is via the equivalent porous media approach (Ghasemizadeh et al. 2012).

Durney et al. (2014) suggest that when considered as an equivalent porous medium, the bulk scale hydraulic conductivity of the study area ranges from less than one, to several tens of metres per day and is generally in the mid-twenties. Along the same line of reasoning, analysis detailed by Golder (2017) suggests that near the MAR trial site, specific yield ranges from 0.05 to 0.2 (Golder 2017). Again, when viewed at the bulk scale as an equivalent porous medium, storativities appear to be in the range of 10^{-2} to 10^{-5} (Durney et al. 2014).

3.5.1 Aquifer properties

ECAN maintains an extensive database of aquifer test parameters and raw data that was used to help conceptualise and guide the modelling and uncertainty analysis. Table 3-1 and Table 3-2 summarise the aquifer test parameters. Well performance tests generally produce lower transmissivity estimates, as the analysis is based on the drawdown in the pumping well. Drawdown in pumping wells is subject to head loss dependent on the efficiency of the well. The inefficiency of a well may be caused by several factors, such as well skin friction or insufficient well development. The overall effect of inefficiency is to lead to additional drawdown, limited to inside the pumping well. Without inclusion of well inefficiency, the actual aquifer drawdown is less. This means aquifer tests with observation wells usually result in higher estimates of transmissivities (Kruseman and De Ridder 2000).

Table 3-1 Aquifer properties from constant discharge tests in the study area (n = 51)

Statistical measure	*Transmissivity (m^2/d)*[7]	*Storativity*
average	1924	0.0011
minimum	101	0.000036
5th percentile	209	0.000041
95th percentile	5300	0.0016
maximum	7500	0.03

[7] Hydraulic conductivity may be derived from Transmissivity where the aquifer thickness is known.

Table 3-2 Aquifer properties estimated from well performance tests (n = 115)

Statistical measure	*Transmissivity (m^2/d)*
average	216
minimum	11
5th percentile	29
95th percentile	450
maximum	1150

3.5.2 Groundwater levels

Early observation data is confounded by two pretrial infiltration pond fillings, one on the 16/5/2016 and another on the 3/6/2016. Despite this, it was still possible to determine which sites clearly responded to the trial which began on the 10/6/2019. Analysis of the water levels recorded as part of the trial showed that only five sites demonstrate a clear response to the recharge water (Figure 3-6). A further eight sites provide a useful record of water levels but failed to show a definitive response to the trial (Figure 3-7). Water level changes recorded at these sites after March 2017 may be in response to the MAR trial. However, quantification of this was confounded by rainfall recharge, with rainfall during this period and the rest of 2017 largely reversing the declines in groundwater levels that had occurred in previous years. The remaining nine sites proved to not be useful for either characterisation of the aquifer or for measuring a response to the trial. Seven sites were excluded from the analysis because they either were dry for large periods of the trial or monitoring began after the beginning of the trial.

Of the five sites that responded to the trial, four showed a similar mounding response of approximately 4 m – 5 m head rise. The fifth site BY20/0152, approximately 1000 m from the MAR site, showed a dramatic 19 m mounding response to the trial. However, this well record appeared to be anomalous when compared to the nearer observation wells. One possible explanation for the large mounding at this well, is leakage down the outside of the casing from a perched overlying layer. At the time the well was developed, it appeared the screen was largely blocked and not responding to the actual part of the aquifer it was screened in, supporting the hypothesis of leakage down the well casing causing the anomalous mounding response.

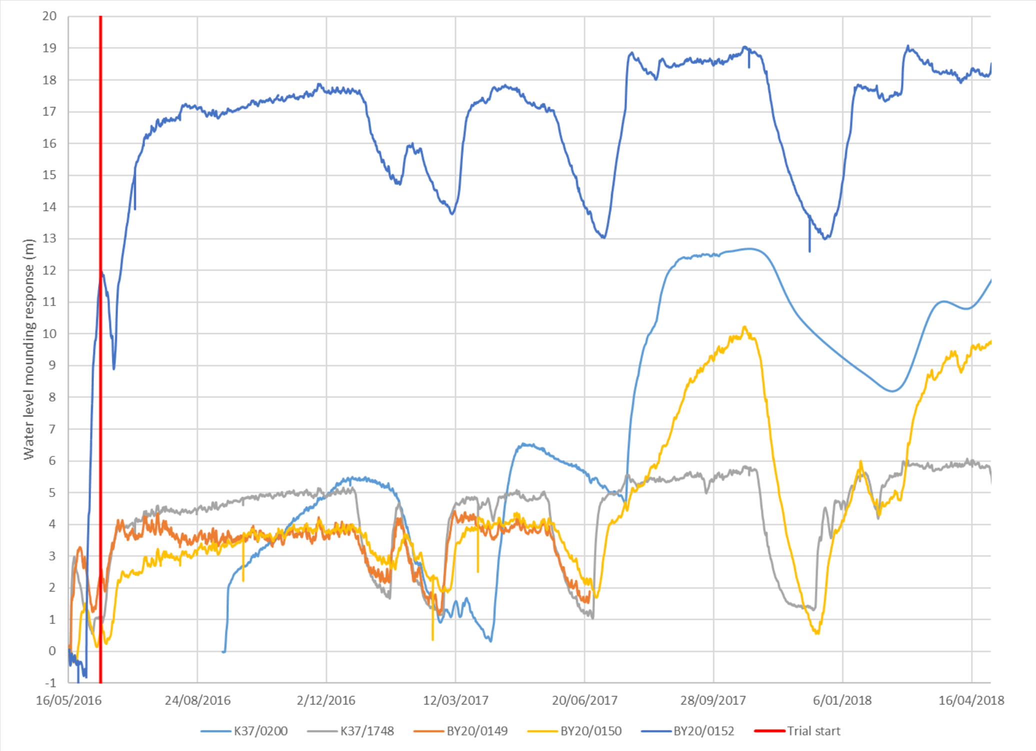

Figure 3-6 Sites showing a clear response to the MAR trial. Initial testing started on 16/5/2016 while the official trial started on 10/6/2016.

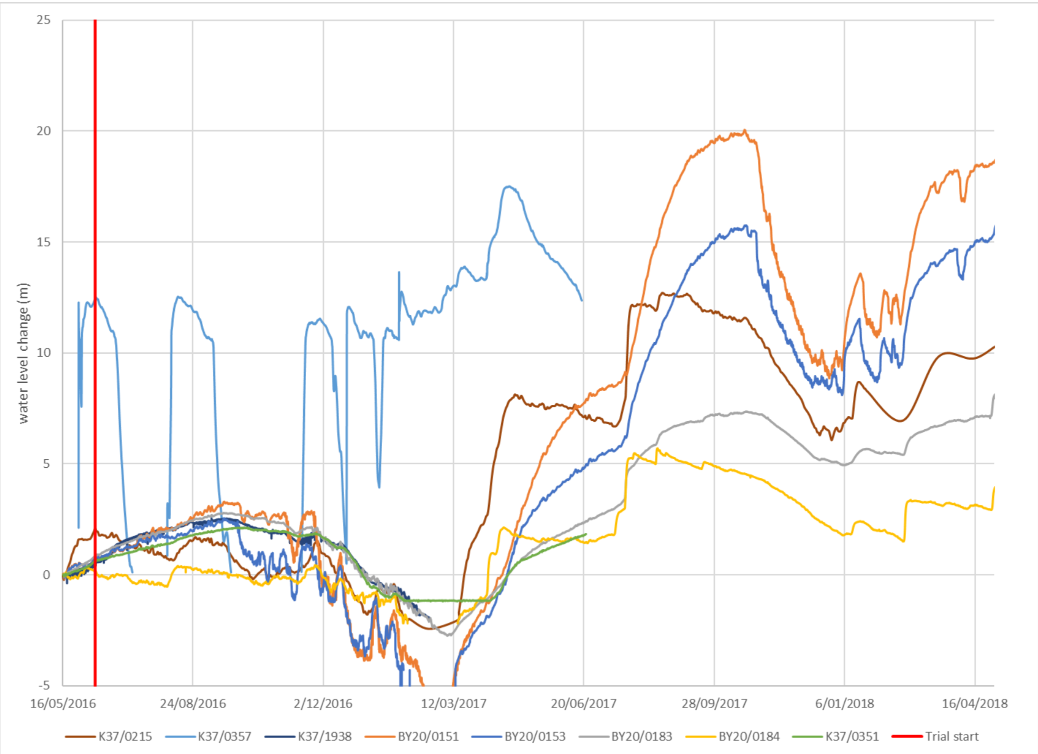

Figure 3-7 Sites with no clear response to the MAR trial, which started on 10/6/2016. Water level fluctuations in these wells are most likely in response to rainfall and pumping.

3.6 Water quality

Simultaneous with the low groundwater levels and declining discharge to the spring-fed water bodies has been rising groundwater concentrations of NO_3-N (Scott 2013). Studies by Hanson and Abraham (2010), who analysed ^{18}O data, suggest those areas with the most degraded water quality are those whose water is sourced from soil infiltration through the intensive pasture areas. The areas that showed the lowest concentrations of NO_3-N coincided with areas where the Ashburton River/Hakatere was actively recharging.

3.7 Land use

The land use in the study area presently consists of pastoral and crop farming, both of which are predominantly irrigated. Two irrigation schemes operate inside the study area (Valetta covering 7000 ha and Eiffelton covering 2700 ha). The Valetta water is supplied from outside the study area via the Rangitata Diversion Race (RDR), which runs across the top of the catchment, while Eiffelton is sourced from local spring and groundwater. There is also extensive groundwater and surface water irrigation that is independently managed. Historically, the irrigation that occurred in the catchment was border dyke irrigation, though in the last decade this has been almost entirely converted to spray irrigation (Durney and Ritson 2014; Durney et al. 2014).

3.8 Groundwater abstraction

For this study, groundwater abstraction is assumed to be approximately 50% (103 million m^3/yr) of the consented volume inside the study area. Water usage data was provided by ECAN. The data show usage ranges from between 30% and 70% in any given year. For example, from 2015 to 2016, water use in Lagmhor was 49% of the allocated volume. As the trial took place at the end of the 2015-2016 irrigation season, the usage from Lagmhor region was adopted as the percentage used for the model area. As the MAR Pilot trial extends across both the irrigation and non-irrigation season, it is essential to know both the location and volume of water abstracted to analyse the result of the trial successfully. The research uses an ECAN database of all consented abstraction takes greater than 5 l/s. Takes of less than 5 l/s are limited to domestic and stock water supplies that have a 10 m^3/day limit for properties less than 20 ha and 100 m^3/day for larger properties. These small takes are ignored as they represent a small part of the total water budget.

3.9 Hinds MAR trial

As previously discussed, MAR is being trialled in the Hinds to assess its effectiveness as a mechanism to redress the over-allocation of water.

Inflows to the trial site were recorded at two electronic recorder stations located in the distribution channel at Flume 1 and Flume 2 (Figure 3-8).

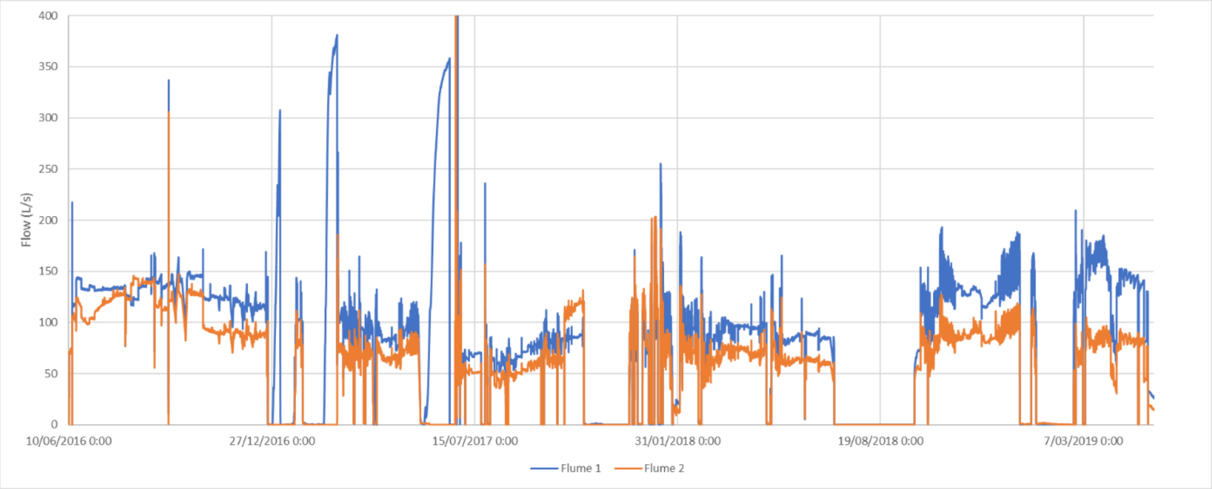

Figure 3-8 Flume flows.

3.10 Groundwater recharge

Average groundwater recharge rates were estimated for the first year of the trial using a soil moisture balance model. In that model, groundwater recharge occurs when rainfall or irrigation exceeds the soil moisture-holding capacity. The recharge over the study area was estimated to be 148.3 million m^3/yr. Figure 3-9 shows the spatial pattern of modelled groundwater recharge. Details of the recharge modelling are provided in Appendix A.

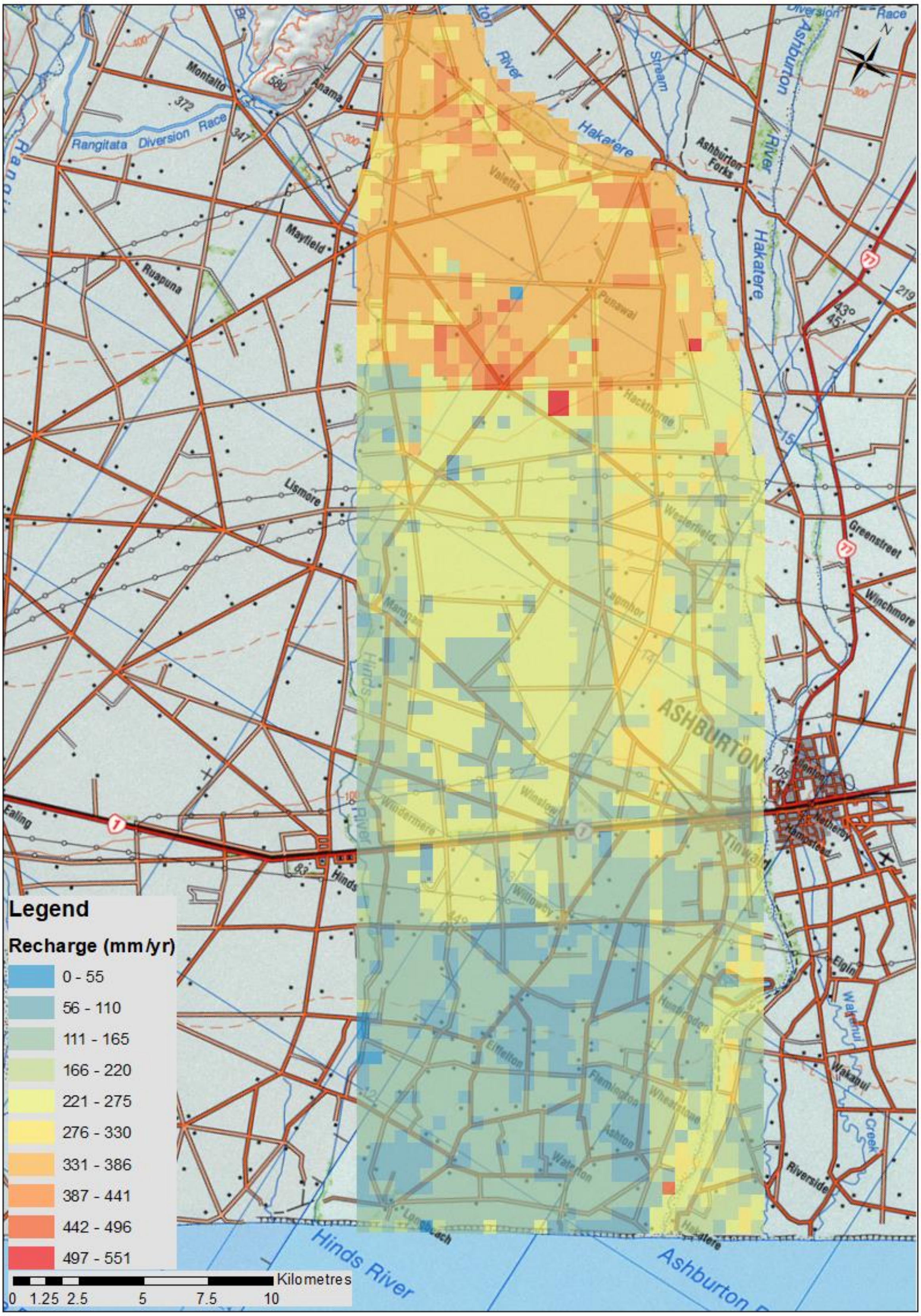

Figure 3-9 Average groundwater recharge across the study area.

3.11 Catchment water balance

Ultimately the conceptual model can be summarised using a water balance. Water balances are useful ways of accounting for the inputs (sources) and outputs (sinks) from the groundwater system. Inputs include rainfall and irrigation recharge through soils, river losses and flow lost from stock water and irrigation races. Outputs include evapotranspiration, groundwater abstraction, discharge to rivers and springs and offshore flow. The estimated groundwater balance for the study area is shown in Table 3-3.

Table 3-3 Study area water balance

Source/Sink	Inputs (million m^3/yr)	Outputs (million m^3/yr)
Ashburton South Branch above Valetta Bridge recorder (assuming loss of 0.5 to 1.5 m^3/s)	*15 - 47*	
Ashburton South Branch below Valetta Bridge recorder (Assumes all gains in flow for the Ashburton River/Hakatere are from Bowyers and Taylors streams and not groundwater)		0
Hinds North Branch (400 l/s)	12.6	
Hinds Main Stem below SH 1 (assuming 50% from outside study area)		9.5
Stock water race losses (assuming 0.5 to 1 m^3/s)	15- 31	
Groundwater recharge (Rainfall and irrigation minus evapotranspiration)	148.3	
Groundwater pumping		103 (assuming 50% of consented allocation used on average)
Spring-fed stream discharge below SH 1		30.6 to 60.6
Off-shore flow (residual)		47.8 to 81.8
Total	190.9 to 254.9 (523,000 to 695,000 m^3/day)	190.9 to 254.9

4 Mathematical modelling

Mathematical modelling of groundwater resources usually takes one of two forms, these being either analytical or numerical in approach.

Through a series of simplifying assumptions, analytical models are usually reduced to closed-form equations that can be solved directly, leading to a discrete although often uncertain answer. Hydrogeology makes use of many analytical equations. The model applications range from simple tasks, such as analysing pumping induced drawdown in wells to more complicated regional modelling using eigenmodels (Bidwell 2005).

Numerical models, on the other hand, are implemented when closed-form analytical solutions cannot be obtained. Instead of producing an exact solution, the numerical model produces an approximation, where the governing equations are approximated at discrete points within the model domain, and the linear system of equations is solved iteratively. Like analytical models, the results are usually subject to high levels of uncertainty. Numerical models are used for more complicated studies and produce results at each grid or mesh cell, interpolating between target observation points.

The choice of an appropriate mathematical model can only be made based on the conceptual understanding of the system. Despite this, in understanding a water resource question, usually, the first step after the development of a conceptual model is to utilise an analytical solution to either approximate the observed results or estimate likely results. To this end, the pretrial data and the MAR trial results were analysed to estimate aquifer parameters and the likely long-term mounding from the MAR trial.

4.1 Analytical modelling

Analytical modelling was used here to:

1. analyse the mounding response in the groundwater observation wells as an inverse constant discharge test to estimate bulk aquifer parameters, such as transmissivity, storativity, specific yield and leakage coefficient. As part of this process, all the observations were corrected for barometric efficiency (removing the water level fluctuations caused by atmospheric pressure).

2. estimate the steady-state mounding response using traditional mounding models, such as the Hantush equation (Carleton 2010). Using both the parameters estimated from the mounding response and text book values.

4.1.1 Analytical estimation of aquifer parameters from MAR trial data

Groundwater mounding caused by the MAR trial has been used to constrain aquifer properties in the surrounding area and to inform analytical Hantush modelling. Analytical solutions, such as the Theis equation (Theis 1935), are commonly used to estimate aquifer parameters from pumping drawdown observation data. Through the inversion of the mounding response (i.e. to recharge), it is possible to use the same solutions to estimate the aquifer properties of the area immediately surrounding the MAR trial site.

The wells that showed an apparent response to the trial are listed in Table 4-1. The water level responses in these wells have been analysed to estimate aquifer properties in the area surrounding the trial.

Table 4-1 Observation well responses

Observed mounding response	*No mounding response*	*No useful data*
K37/0200	K37/0215	K37/0204
K37/1748	K37/0351	K37/1749
BY20/0149	K37/0357	K37/0972
BY20/0150	K37/1938	K37/3050
BY20/0152	BY20/0151	K37/1540
	BY20/0153	K37/0383
	BY20/0183	K37/2538
	BY20/0184	

The data preparation for analytical analysis requires water level corrections for barometric efficiency. Barometric efficiency relates to the water level fluctuations in wells caused by atmospheric pressure. These fluctuations need to be removed from the observation data to enable quantification of water level changes caused by pumping or recharge.

Assessment of barometric efficiency can also help inform the conceptualisation of the aquifer and selection of the appropriate analytical solution. Wells screened in unconfined conditions have between 80% and 100% barometric efficiency, meaning their water levels respond directly to atmospheric changes. This contrasts with fully confined aquifers, where the water levels generally show less response to barometric pressure changes (25% to 75% efficiency) (Solinst 2019). Barometric efficiency analysis is presented in Appendix B, with the efficiency of the wells close to 100% supporting an unconfined conceptual model.

The analytic solution selected for the analysis is dependent on the conceptual aquifer setting. Initial analysis was conducted assuming unconfined conditions and used both the Cooper-Jacob solution and the Theis solution (analysis is presented in Appendix B). As neither solution could adequately explain the responses observed, a third analytical solution, the Neuman-Witherspoon (1969) two-layer aquifer solution was used to model the mounding. All the solutions employed required high estimated storativity/specific yield to match the length of time it took for the observation wells to respond to the recharge. However, further data analysis suggested the delayed mounding response was due to vadose zone influences rather than aquifer properties. Table 4-2 shows the estimated parameter range for transmissivity and storativity/specific yield.

Table 4-2 Estimates of aquifer parameter range from the analytical analysis of MAR trial mounding

	Transmissivity (m^2/day)	*Specific yield*
Upper bound	800	0.4
Lower bound	200	0.02

While the analytical analysis has not been able to fully estimate the specific yield responses, it proved useful for estimating transmissivity and provided a range of possible specific yields that could be employed to constrain stochastically generated Hantush (1967) mounding assessments. The Hantush (1967) solution for mounding beneath a rectangular recharge basin is given in Equation 1.

Equation 1

$$h^2 - h_i^2 = (\frac{w}{2K})(vt)\{S_y * \left(\frac{l+x}{\sqrt{4vt}}, \frac{a+y}{\sqrt{4vt}}\right) + S_y * \left(\frac{l+x}{\sqrt{4vt}}, \frac{a-y}{\sqrt{4vt}}\right) + S_y * \left(\frac{l-x}{\sqrt{4vt}}, \frac{a+y}{\sqrt{4vt}}\right) + S_y * \left(\frac{l-x}{\sqrt{4vt}}, \frac{a-y}{\sqrt{4vt}}\right)\}$$

where:

- h is the head beneath the mound (m)
- h_i is the static head before recharge (i.e., the initial saturated thickness of aquifer) (m)
- w is the recharge rate (m)
- K is the hydraulic conductivity of the aquifer (m/day)
- v is the diffusivity, where v=Kb/Sy (m^2/day)
- b is the average aquifer thickness (m)
- S_y is the specific yield of the aquifer [dimensionless]
- t is the time elapsed since recharge began (days)
- l is ½ the recharge basin length (m)
- a is ½ the recharge basin width (m)
- x is the distance from the centre of the recharge basin in the x-direction (m)
- y is the distance from the centre of the recharge basin in the y-direction (m)

Two approaches to the analytical estimation of mounding were investigated. One, a stochastic ensemble of models (using a spreadsheet model modified from Poeter and McCray, 2008) were built, that produced mounding estimates at four distances from the recharge basin. These results were used to produce the 5th, 50th and 95th percentiles of mounding that may be anticipated from constrained textbook parameter values. The underlying parameters used in mounding assessments were:

- h_i = 10 *m*
- *K* = constrained to between 1 *m/d* to 100 *m/d*
- *Sy* = constrained to between 0.01 to 0.3
- *w* = 9000 m^3/d
- *a* = 50 *m*
- *l* = 50 *m*

In the second approach, mounding estimates were produced guided by the parameters estimated from the analytical drawdown solution. For this purpose, the mounding tool in Aqtesolv™ was used. Outputs were in the form of gridded surfaces of mounding in the 10 km surrounding the recharge basin.

4.2 Numerical modelling of the MAR trial

Numerical modelling was carried out concurrently with analytical modelling. Two numerical models were developed using MODFLOW-NWT (Niswonger et al. 2011); one, a homogeneous parameter model and the other a more complicated, highly parameterised (involving a large number of model parameters) model using pilot points. MODFLOW–NWT was developed as a separate numerical formulation of the MODFLOW 2005 code that is designed to deal with drying cells. MODFLOW modelling was conducted in Groundwater Vistas.

Numerical modelling was used in this study to focus on the range of possible mounding outcomes of the trial, and to provide an estimate of the confidence in the model outputs. To this end, a stochastic modelling approach was adopted, where Monte Carlo techniques were used for uncertainty estimation.

Several steps were involved in numerical modelling:

1. Development of a recharge time series.
2. Transformation of the conceptual model into a MODFLOW numerical structure with grid refinement around the MAR Pilot area, introducing groundwater level and river flow observations held by ECAN as model calibration targets.
3. Calibration of the MODFLOW model as a steady-state flow model using the Pilot Point PEST (Doherty 2003) method.
4. Uncertainty analysis of the flow model results using Monte Carlo techniques. Specific measures considered when conducting the uncertainty analysis were:
    - assessment of how far downgradient the 0.1 m of mounding propagates

- assessment of head change at a set location
- area of head change greater than 0.1 m

6. Development of a transport model using MT3DMS to investigate MAR plume propagation
7. Use the model to predict the mounding response at the end of the trial and estimate the associated uncertainty.

4.2.1 Model domain and grid

The model domain extended from just below the Rangitata Diversion Race Management Ltd main race at its inland extent to the coast and from between the Ashburton River/Hakatere in the north and the Hinds River/Hekeao in the south. The model was discretised on a 500 m by 500 m grid, with telescoping to 100 m by 100 m resolution in the immediate trial site vicinity (Figure 4-1).

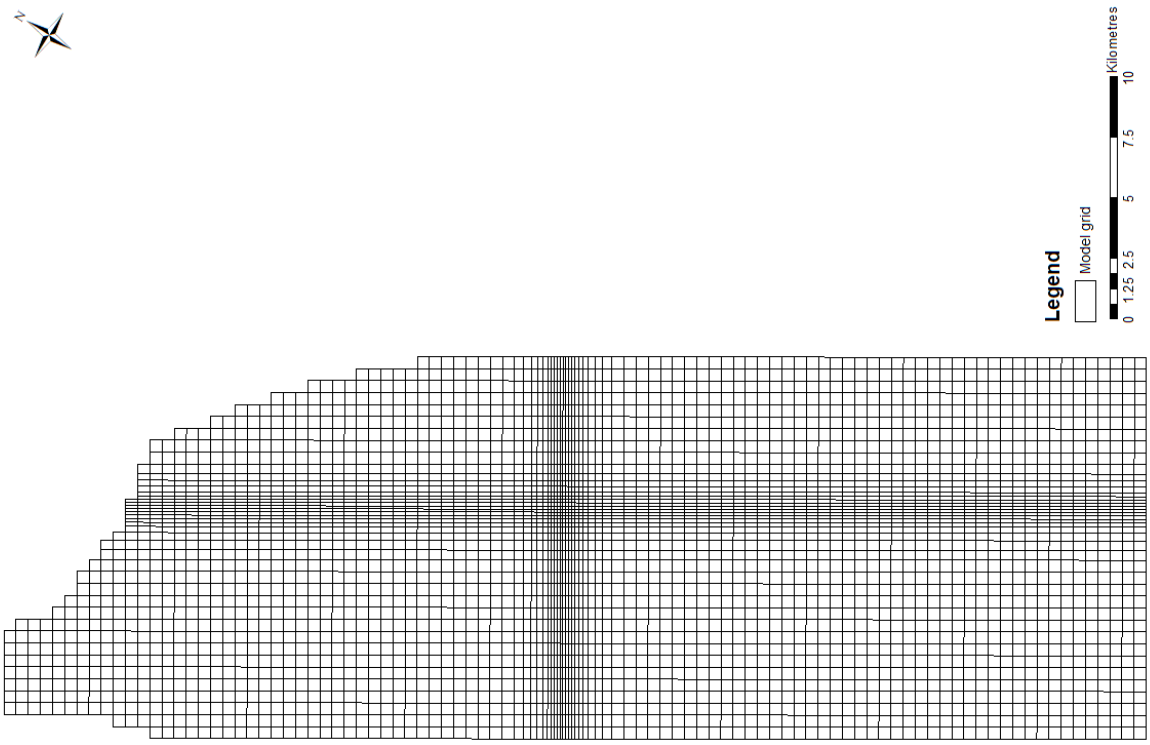

Figure 4-1 Model grid discretisation.

4.2.1 Model layers and thicknesses

Initially, an arbitrary 10 layers were used in the discretisation of the model. However, preliminary manual calibration showed that this spatial resolution was insufficient to capture the near-surface aquifer conditions in adequate detail. Subsequently, an additional 3 layers were added to the first 100 m below ground surface. Layer 1 extends from the topographic surface to between 10 m and 70 m below surface. The layers vary in thickness from several metres at the coast, to up to 75 m at the inland boundary. Appendix C shows the adopted model layer elevations.

4.2.2 Head dependent boundary conditions

Head dependent boundary locations are shown in Figure 4-2. The inland boundary conditions were specified as a general head boundary of 420 m elevation, to represent the water table at the inland extent of the aquifer located at the base of the foothills. A general head boundary was used, as conceptually some volume of flow must enter the model domain through the inland extent. However, given the small physical extent before the base of the foothills, the daily volume of water available to enter through this model boundary was small. Therefore, the conductance of the boundary condition was adjusted in order to allow some inflow, while also preventing large rates of water entering the model domain; thereby maintaining the conceptual water balance. Initially, the coastal boundary used a saltwater density adjusted specified head for the centroid of each layer (Motz and Sedighi 2009). Preliminary manual calibration revealed that for the deeper layers, this could not explain the groundwater elevations in coastal monitoring wells. After some experimentation, layers 9 to 13 were set to no flow boundaries, which produced elevations closer to observations. Density adjusted elevations for layers 1 to 8 are shown in Table 4-3. All other boundary extents were set as no flow.

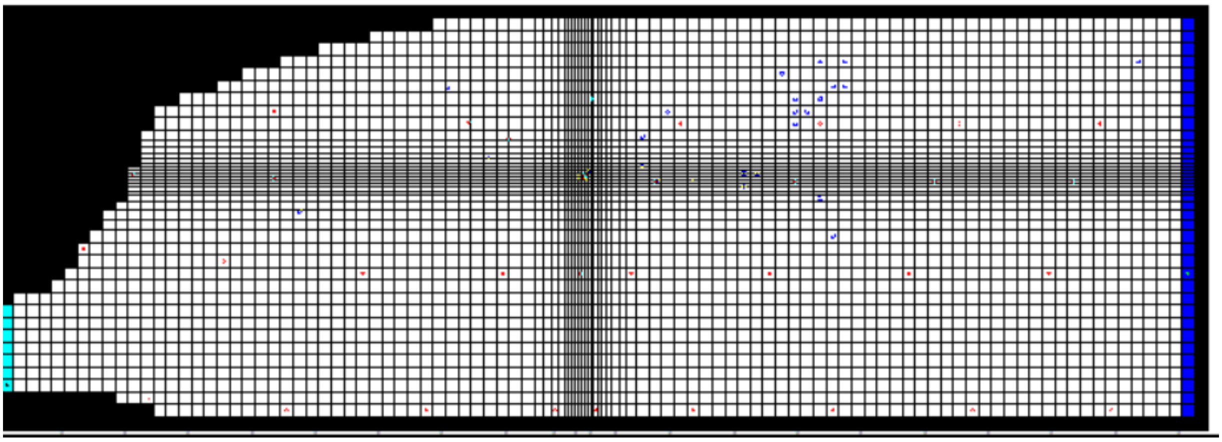

Figure 4-2 Model boundaries - light blue = general head; dark blue = specified head, blue dots are pumping wells while red dots are pilot point locations.

Table 4-3 Density adjusted specified head boundary elevations

Layer no.	*Boundary elevation m (0 + 0.125 x elevation cell centroid)*
1	0.11
2	0.20
3	0.26
4	0.34
5	0.46
6	0.65
7	0.81
8	0.94

4.2.3 Unsaturated zone

The model does not explicitly represent the unsaturated (vadose) zone. Instead, unsaturated flow (recharge) is represented using a land-surface recharge layer. All recharge is applied directly to the saturated zone.

4.2.4 Saturated zone properties

The behaviour of water flow through an aquifer is described by Darcy's law (Equation 2).

Equation 2

$$Q = K.i.A$$

Where:

- Q is discharge (*unit volume per unit time*)
- K is hydraulic conductivity (*unit length per unit time*)
- i is hydraulic gradient (*unitless*)
- A is the area of discharge (*unit area*) (Freeze and Cherry 1979).

Numerical models are, put very simply, a framework for the application of Darcy's law and conservation of mass. As such, inside a steady-state numerical groundwater model, the primary aquifer property of concern is hydraulic conductivity. Hydraulic conductivity controls the maximum rate of transmission of water through the aquifer and is responsible, along with recharge sources, for the elevation of the water table within the aquifer. It is defined as the volume of fluid at the existing kinematic viscosity, that will move in a unit time under unit hydraulic gradient, through a unit area measured at right angles to the direction of flow (Freeze and Cherry 1979).

When modelling the time-varying behaviour of an aquifer, storage properties also need to be considered (i.e. the specific yield for unconfined aquifers and specific storage for confined aquifers). As the numerical model developed in this study is steady-state, storage properties are not considered further.

Initial aquifer properties were set based on analysis of pump testing data held by ECAN, as presented in Table 3-1. The uniform homogeneous layer model used the median estimated hydraulic conductivity for its initial parameters. In contrast, the initial hydraulic conductivities used in the heterogeneous modelling used a combination of values based on aquifer test parameters at the spatial location where they intercepted the model grid in 3-D, and the homogeneous models parameter estimates where aquifer test data was lacking. Transmissivity is converted to hydraulic conductivity based on the thickness of the aquifer. In the Hinds Plains, there is no clear distinction between aquifers. One approach is to assume the entire saturated zone is one aquifer and to use this thickness. However, variographic analysis of aquifer test data and specific capacity in other parts of Canterbury suggests that at most, the aquifer test data relates to an aquifer thickness of 30 m. With no clear way to distinguish the thickness of the aquifer beneath the study area, the thickness used in estimating initial hydraulic conductivities was assigned based on the model layer thickness for the cell, inside which each test lies (Table 4-4). Automated calibration was free to vary the initial conditions as much as needed to achieve calibration.

Table 4-4 Hydraulic conductivities estimated from Environment Canterbury held aquifer test data

Statistic	*Hydraulic conductivity (m^2/d)*
Min	1
5th percentile	4
Median	52
Mean	97
95th percentile	322
Max	539

4.2.5 Rivers gains and losses

The Ashburton/Hakatere and Hinds/Hekeao Rivers are represented using the MODFLOW Streams Package (STR) (Figure 4-3). This package has stage-dependent loss or gains controlled by a combination of the water table elevation, specified inflow from the upstream boundary, underlying aquifer properties and a streambed conductance factor. Inside the model, observed river flow losses and gains are represented as calibration targets that are met by varying the streambed conductance and hydraulic conductivity of the aquifer.

The stock water supply network is also represented in the model using the STR package, with a calibration target of approximately 86,400 m^3/day loss to groundwater. The STR package was chosen because it allows the total exchange with the saturated zone to be controlled by the rate of water entering the surface water network, thereby avoiding conceptual mass balance errors. Unlike more simple representations of river features, the streambed conductance is controlled by the width of the stream and thickness of the streambed material. Streambed conductance is a calibration parameter, while the width of the stream was specified based on aerial photograph interpretation and thickness arbitrarily set to one metre.

The widely distributed network of coastal drains is represented in the model using the drains package (Figure 4-3). The drains package acts similarly to the MODFLOW rivers package, except the flux is only one way (out of the model). The controlling factors are water table elevation, drain elevation level and a conductance factor. Drain elevation level was based on analysis of ECANs LiDAR collection and set to 1 m below the model topography. The conductance factor is a calibration parameter, while the water table elevation is calculated by the model.

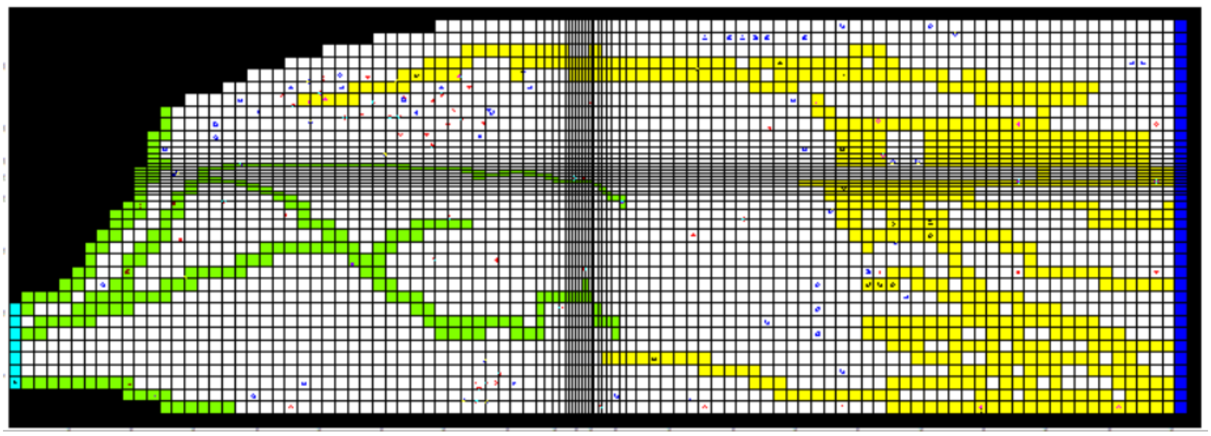

Figure 4-3 Spatial representation of the streams (green) and drain package (yellow).

4.3 Parameter sensitivity and model calibration approach

4.3.1 Sensitivity analysis

Sensitivity analysis investigates the sensitivity of model results to the various parameters used in the simulation. It is an integral part of the calibration process, identifying which parameters have the greatest influence on the calibration fit. In highly parameterised models, it is even more critical, as identifying the most sensitive parameters enables these to be focused on during the calibration process, thereby increasing the computational efficiency. PEST has several tools that tackle sensitivity analysis, such as SVD. SVD achieves sensitivity analysis by running a simulation for each parameter. First, a simulation is run using the initial parameter values; then an individual parameter is varied and the simulation re-run. The simulation statistics are compared to those generated before the parameter was varied. This process is repeated for each parameter (Doherty 2019).

4.3.2 Calibration

Calibration is the process of adjusting model parameters within reasonable bounds so that simulated outputs resemble actual observations. Calibration can be manual or automated. Manual calibration involves the manual adjustment of calibration parameters. The key benefit of the manual calibration is that it teaches the modeller about the behaviour of the aquifer system being modelled. Also, it is useful at an early stage for ironing out bugs in the model that may otherwise be masked by the automated calibration process.

Automated calibration largely removes the requirements for the continual intervention of the modeller and can lead to more statistically robust results. Further, the automated calibration process often produces better fits to observation data. The usual statistic applied for the determination of calibration fit is the root mean square error (RMSE). RMSE describes the average error of the simulated value to observation (simulation minus observation).

4.3.3 Uniform homogeneous parameters

The simplest numerical modelling approaches use lumped homogeneous parameters. The homogeneous approach is computationally inexpensive because few parameters are being estimated. Thus, the approach is computationally fast and can provide useful answers in short timeframes. Further, the use of homogeneous parameters and parameter zones most closely follow the assumptions used in analytical models.

Due to the simplification of reality required for homogeneous models, the model predictions are more approximate than highly parameterised models. Though in and of itself, this does not mean the model is a worse predictor, only that history matching of previous events is less likely than with a highly parameterised heterogeneous model.

The homogeneous modelling conducted for this research used the steady-state approach; the only parameters estimated were hydraulic conductivity (horizontal and vertical) and streambed conductance.

4.3.4 PEST Pilot Points

The alternative approach to using homogeneous or zone parameters is to estimate the properties of every single grid cell. Adjusting every cell is impractical in terms of computational efficiency and lacks conceptual basis. Because the estimated properties of the aquifer are based on observation points, it is only possible to estimate the average properties between, or at the observation points and not at the individual model cells that are generally far more abundant than observations. A compromise between homogeneous zones and estimating properties at every cell is achieved through the pilot point technique. Pilot points act as parameter surrogates and two, though not mutually exclusive, placement approaches may be adopted. One where pilot points are placed in a regular gridded network and the other where they are placed irregularly, ideally either at the observation points when estimating porosity, or else midway between the observation points in the direction of flow when calculating hydraulic conductivity (Doherty et al. 2010a; Doherty 2019). These surrogate parameters can then be estimated using inverse modelling techniques. A continuous surface can then be created of the properties generated by interpolating between the pilot points. This continuous surface is then introduced into the modelling software with a property value at each grid cell. The technique allows for very detailed approximations of aquifer properties that smoothly transition across the model grid, maintaining both a physical/mathematical basis and numerical efficiency (Doherty 2003).

4.4 Transport modelling

Improved water quality is one of the stated goals of the Hinds MAR trial (Golder 2017). MT3DMS (Zheng and Wang 1999) has been employed to model advective transport of the plume of clean MAR water. MT3DMS stands for Modular Transport 3D Multi-Species model. MT3DMS is designed to model multi-species reactive transport in a groundwater flow field modelled using MODFLOW (Zheng and Wang 1999).

For the Hinds MAR trial, the key water quality measure is the replacement of groundwater high in NO_3-N with Rangitata River water, with its very low concentrations. To this end, the monitoring programme has focussed to a large extent on NO_3-N (Golder 2017). Modelling has focused on changes in NO_3-N in response to the trial.

Background conditions across the study area for water quality are unknown, and while being extensively sampled at point locations, the study area has no accurate maps of loads of NO_3-N entering the groundwater system. Modelling of provisioned catchment loads have been carried out by ECAN (Scott 2013) in support of the LWRP plan change 2. However, these are hypothetical scenario leaching rates. Therefore, instead of introducing a clean plume into a contaminated aquifer, the MAR trial water has been modelled as a non-reactive contaminant with a concentration of 100 mg/l. Assessment is based on the proportion of water in the model cells that comes from the MAR trial after five years of simulation.

Transport modelling requires knowledge or estimates of porosity and dispersivities. For the study area, neither of these is known, and instead, they have been investigated as part of the modelling process. Model resolution (100 m by 100 m at its finest) is too coarse to allow an accurate estimation of dispersion; further data sparsity reduces the ability to provide meaningful estimates. Initial experimentation at the finest (100 m by 100 m) model resolution demonstrated that if dispersion is less than 10 m, it is unlikely to influence the results significantly (Figure 4-4).

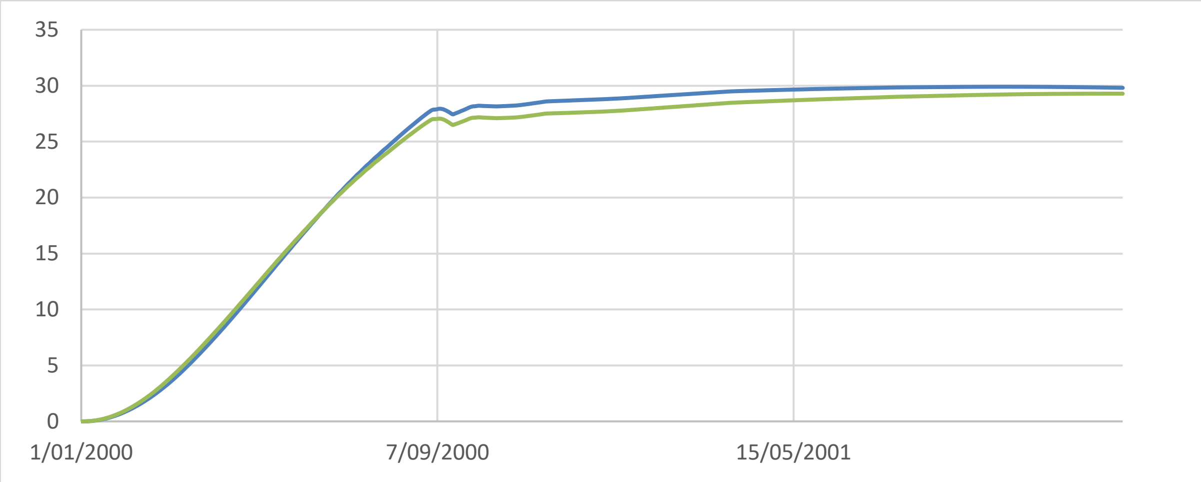

Figure 4-4 Difference in concentration breakthrough between nominal longitudinal (< 1e-4 m) dispersivity and 10 m longitudinal dispersivity at 100 m by 100 m resolution.

Review of literature suggested that at the model scale, longitudinal dispersivities were unlikely to be significantly greater than 10 m (Figure 4-5) (Schulze-Makurch 2005).

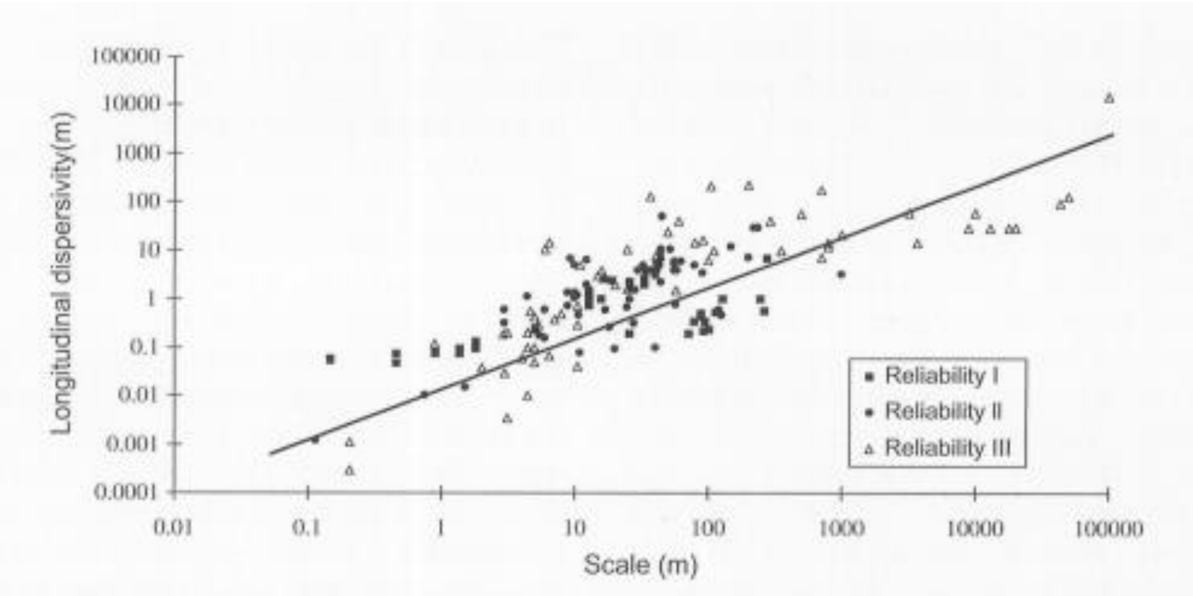

Figure 4-5 Relationship between longitudinal dispersivities and scale from Schulze-Makurch 2005.

Given the lack of data with which to accurately estimate dispersion, a decision was made to adopt nominal of dispersivities of $1x10^{-4}$ m longitudinally, $1x10^{-5}$ m transversely and $1x10^{-6}$ m vertically. Doing so effectively suppresses dispersion and instead allowed focus on the calibration of porosity.

This decision is in line with advice from H Prommer (pers comms. Henning Prommer September 2016). The water quality changes in wells BY20/0152 and K37/0200, which responded to the trial start after approximately three months and eight months respectively (Golder 2017), have been used to estimate the porosity of the aquifer.

5 Results

5.1 Analytical modelling

5.1.1 Stochastic ensemble

The stochastic ensemble generated 3150 separate mounding estimates. Of these, 87 were rejected as they suggested mounding above the ground surface. The remaining results are summarised in Table 5-1 and Figure 5-1, which shows the respective mounding percentiles on a distance basis. The stochastic ensemble suggests there is a less than an 8% probability of the MAR trial producing mounding greater than 0.1 m at 5 km from the trial site.

Table 5-1 Statistics from the stochastic ensemble (individual point) n=3150

Statistic	*Mounding at 100 m*	*Mounding at 500 m*	*Mounding at 1000 m*	*Mounding at 5000 m*
5th percentile	3.87	2.06	0.02	0.00
50th percentile	6.24	3.40	1.71	0.00
Mean	8.03	4.42	2.07	0.06
95th percentile	19.04	12.33	4.73	0.44

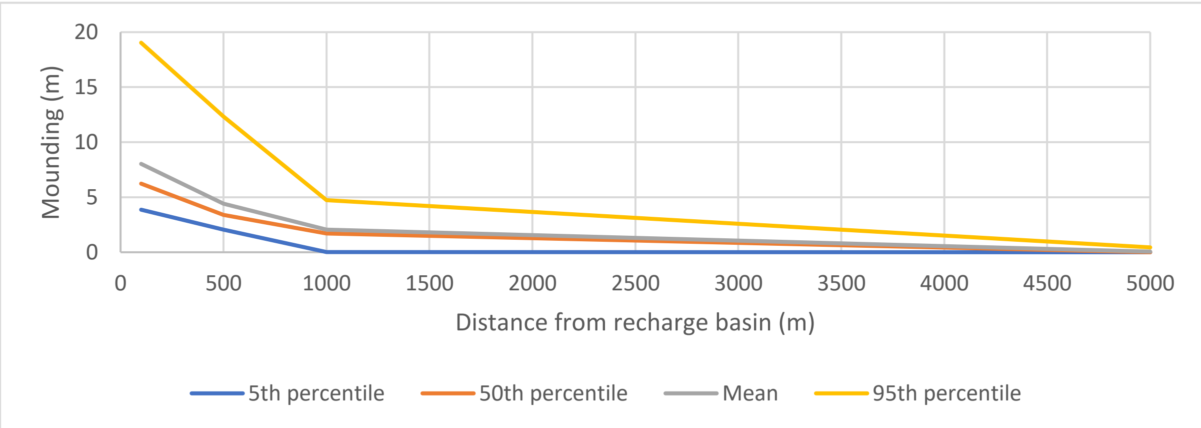

Figure 5-1 Analytical mounding at specific distances from the recharge basin.

5.1.2 Calibration constrained mounding

In contrast with the stochastically generated ensemble of predictions, the calibration constrained predictions suggest that mounding greater than 0.1 m will propagate slightly coastwards of SH1, e.g. greater than 10 km from the recharge basin (Figure 5-2 to 5-4). Though the mounding response within 5 km of the trial site is likely to vary considerably, from upwards of 10 m with the greatest mounding, to less than 6 m with the least mounding.

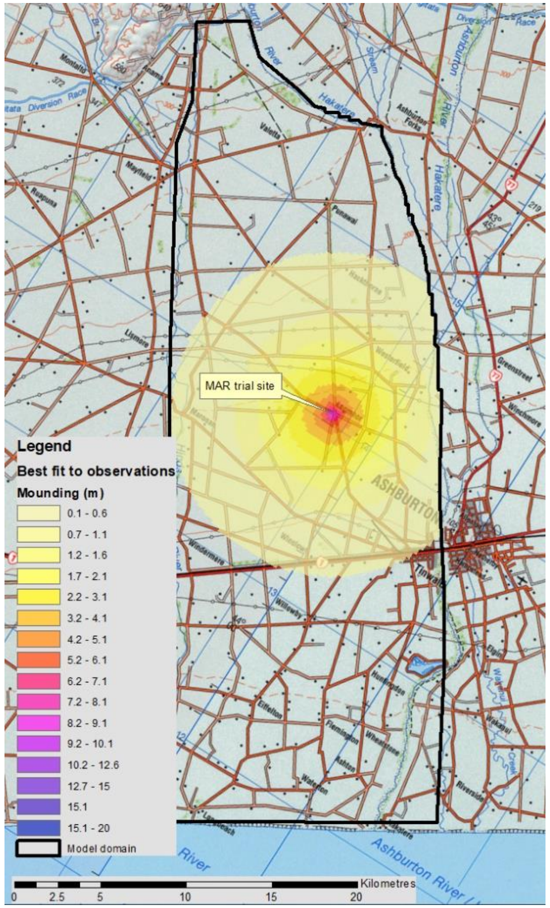

Figure 5-2 Mounding response to the best parameter set from analytical modelling.

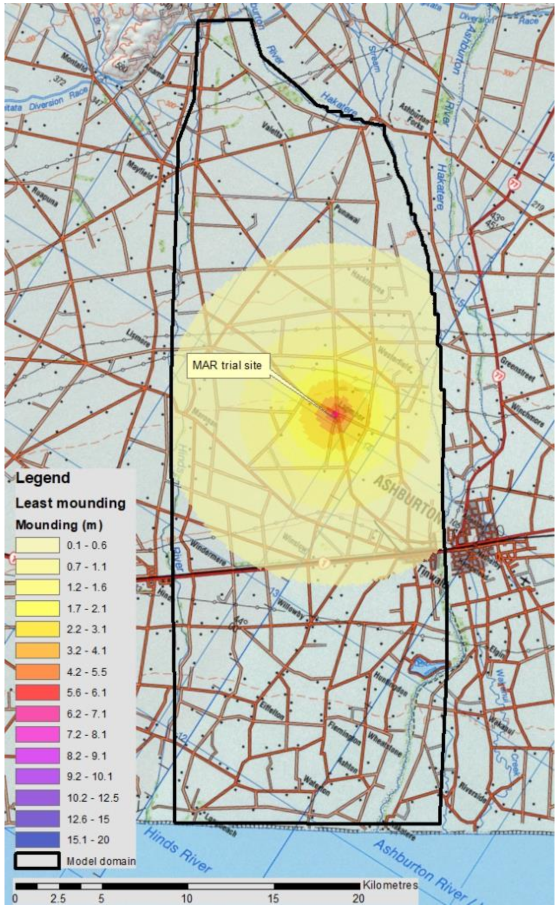

Figure 5-3 Least mounding response to parameter sets from analytical modelling.

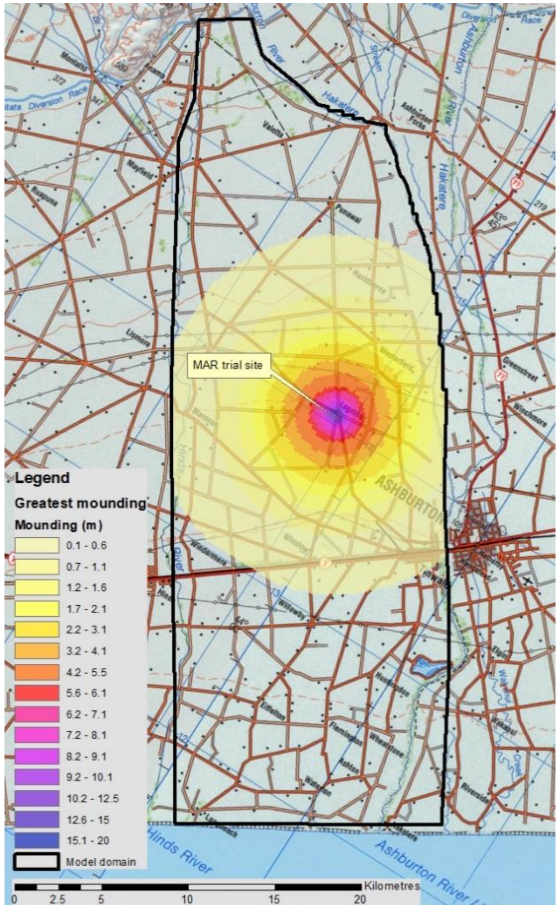

Figure 5-4 Greatest mounding response to parameter sets from analytical modelling.

5.2 Deterministic numerical modelling

As discussed in Chapter 4, two MODFLOW models were developed; the first using homogeneous parameters and second calibrated using pilot points. Both model calibration approaches used the model structure discussed in Chapter 4.

5.2.1 MODFLOW homogeneous modelling

The first numerical model developed of the study area was calibrated using a combination of manual calibration and automated calibration using PEST (Doherty 2003). The model used 98 observation points weighted to the reliability of the data. Reliability was determined from the number of observations at each location and the dates of the observation. Higher weightings were given to sites with multiple observations, with still greater weights given to those with many observations recorded after 1/1/2014. Sites with only a few observations were included to help guide the calibration and prevent the model producing results that were physically unreasonable, i.e. to constrain the water level elevations in areas with no other data.

Uniform horizontal and vertical conductivities were applied across the model domain and varied both manually and automatically using PEST to optimise the model to the observation data. The final adopted hydraulic conductivities were Kh = 11.24 m/day and Kv = 0.0122 m/day. Overall the model performed adequately but failed to provide a match to all observation points (Figure 5-5). Further, the model overpredicted water levels in the area surrounding the trial site. Despite the inability to predict all head elevations, the model mass balance is in line with the estimated catchment water balance (Table 5-2). The inability to accurately represent the spatially distributed groundwater levels suggests the model is under-parameterised, and that representing the aquifer system with homogenous parameters is unrealistic.

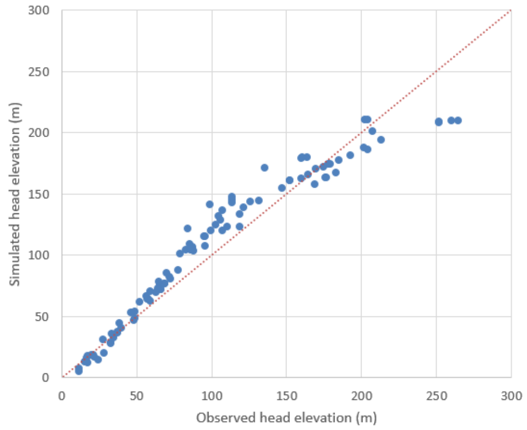

Figure 5-5 Homogeneous model calibration fit with general head boundary.

Table 5-2 Homogeneous model mass balance with general head inland boundary

Mass balance	*Inflow m^3/day*	*Outflow m^3/day*
MAR	*9000*	
Fixed head boundary	0	232671
General head boundary	5895	0
Stream package	231612	14434
Recharge	407148	
Drain package	0	170745
Groundwater abstraction	0	235808
Total	653655	653658
Percent error	-5.00E-04	

Analysis of the MAR induced mounding revealed that the choice of the inland model boundary was having an impact on the model results. The initial model was calibrated using a general head boundary for the inland extent. The effects of this boundary choice had been estimated by transforming the model into a fixed head boundary condition, equivalent to the head elevation estimated using the general head boundary and the model re-run. Calibration was maintained with the changed boundary type (Figure 5-6 and Table 5-3). The calibration statistics were similar with both model boundary types (Table 5-4) and showed a scaled RMSE of less than 6%.

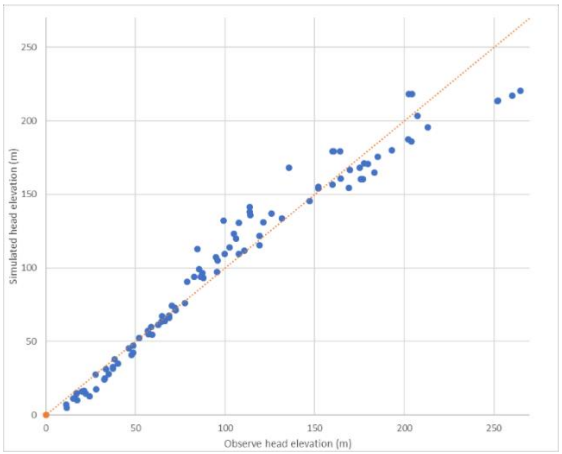

Figure 5-6 Homogeneous model calibration fit (fixed head boundary).

Table 5-3 Homogeneous model mass balance with fixed head inland boundary

Mass balance	*Inflow m^3/day*	*Outflow m^3/day*
MAR	9000	
Fixed head boundary	15110	239825
Stream package	231637	14535
Recharge	405532	
Drain package	0	171115
Groundwater abstraction	0	235808
Total	661279	661283
Percent error	-4.00E-04	

Table 5-4 Homogeneous model calibration statistics

Statistical measure	*Homogeneous model (m) (general head boundary)*	*Homogeneous model (m) (specified head boundary)*
Residual mean	0.41	-1.82
Residual standard deviation	13.76	13.69
Absolute residual mean	9.65	9.44
Residual sum of squares	1.88×10^4	1.89×10^4
RMSE	13.76	13.82
Minimum residual	-32.95	-39.05
maximum residual	44.27	36.95
Range of observations	253.12	253.12
Scaled residual standard deviation	0.054	0.054
Scaled absolute mean	0.038	0.037
Scaled RMSE	0.054	0.055
Number of observations	98	98

Simulated mounding is truncated down-gradient rapidly by the presence of the drainage network, which efficiently conveys away any water that reaches this area (Figure 5-7). The mounding in the deeper model layers does propagate further towards the coast, although the mounding that does so is less than the assessment threshold of 0.1 m. The model results share a commonality in that both designs show that the coastal drains effectively minimise the mounding response in the coastal region of the model. However, as expected, the fixed head inland boundary causes the model to simulate significantly less mounding at the inland extent than the general head boundary allows (Figure 5-8). This difference in results highlights the importance of model boundary conditions and their impacts on model predictions. Both results show that mounding is most pronounced near the recharge basin; however, like the calibration constrained analytical modelling, the mound propagates further than would be suggested by the analytical modelling ensemble.

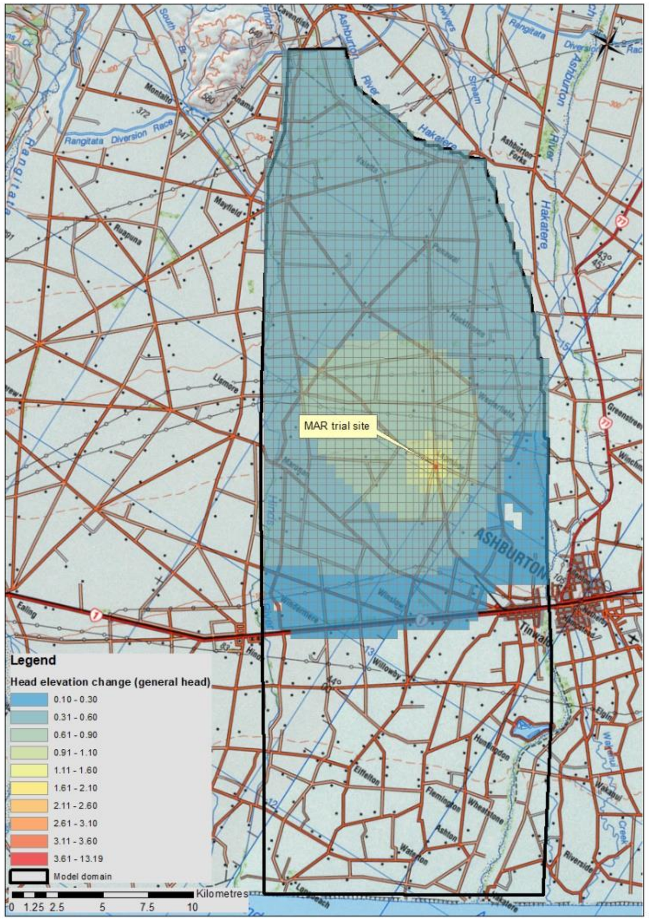

Figure 5-7 Homogeneous model water table mounding with the general head boundary at the inland extent.

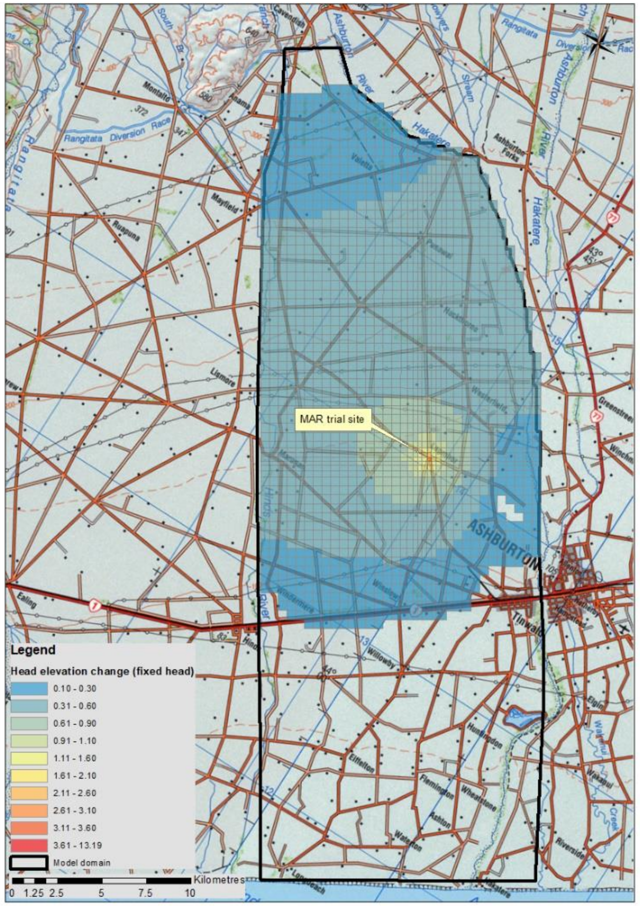

Figure 5-8 Homogeneous model water table mounding with a fixed head boundary at the inland extent.

5.2.2 **Transport modelling**

MT3DMS was utilised to model transport. Nominal dispersivities discussed in Chapter 4 were used along with manual and automated calibration of porosity. Porosities based on the breakthrough of the MAR water at monitoring wells K37/200 and BY20/0152 appeared to be approximately 0.01 to 0.06. Automated PEST calibration resulted in porosity of 0.055. However, manual calibration seemed to fit the observed breakthrough better at well BY20/0152, where automated calibration skewed results towards well K37/0200. Results for the manual calibration after 280 days and 5 years are shown in Figures 5-9 to 5-12.

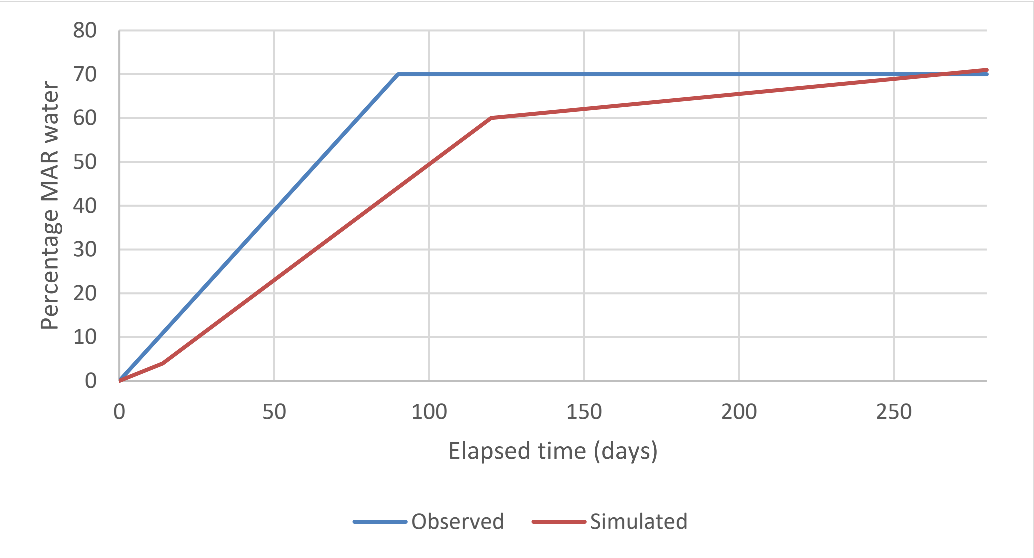

Figure 5-9 Plume breakthrough BY20/0152 (porosity 0.01).

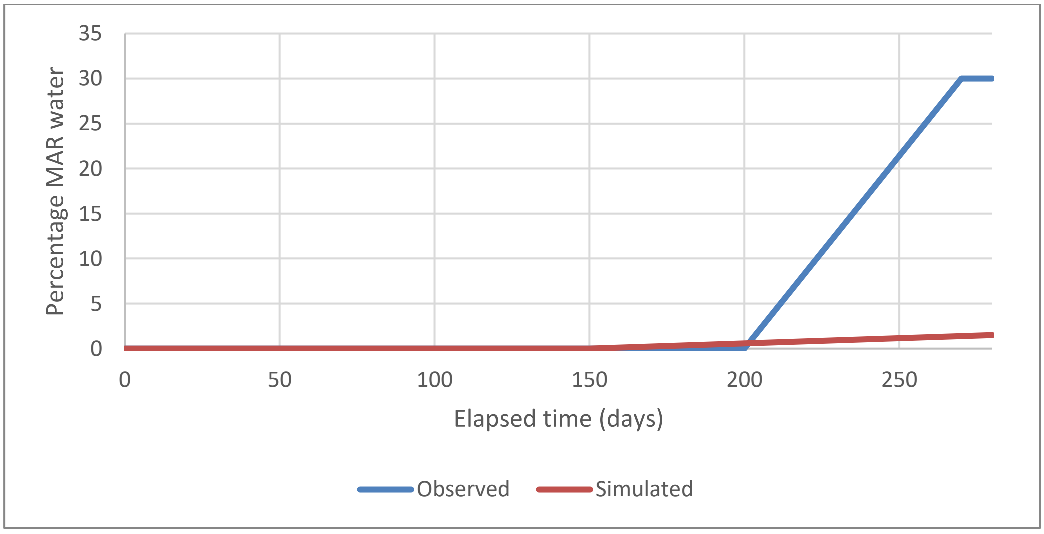

Figure 5-10 Plume breakthrough K37/0200 (porosity 0.01).

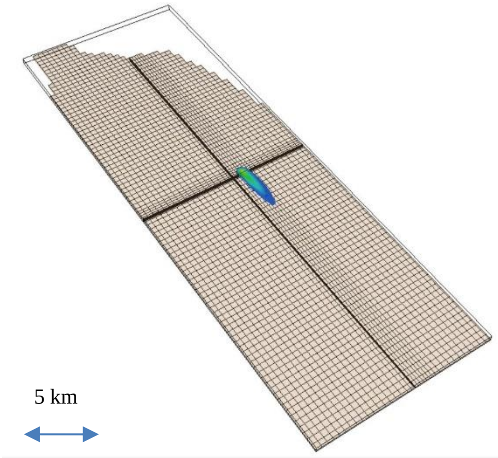

Figure 5-11 Percentage of water from MAR plume after 280 days (porosity 0.01).

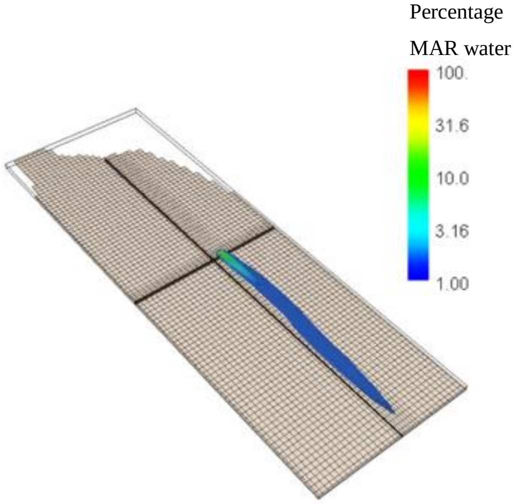

Figure 5-12 Percentage of water from MAR plume after 5 years (porosity 0.01).

5.2.3 MODFLOW Heterogenous modelling

The second calibration approach used highly parameterised spatially varying hydraulic conductivity, adopting the pilot point PEST method. Hydraulic conductivity varied both in the horizontal and the vertical axis. A total of 828 Pilot Points and 5 conductance factors were used for the calibration. SVD assist was employed to accelerate the calibration process and later allowed the use of NSMC uncertainty analysis. Two calibration criteria were used; the first calibration approach used specified head targets (groundwater levels in wells) and flux to or from the streams and drainage network. The second approach introduced concentration targets into the automated calibration process.

5.2.4 Initial calibration using only head elevation and flux targets

The initial calibration only used pretrial observation data. A good fit was achieved to the observation data, however when MAR was introduced into the model using an injection well, many model cells adjacent to the recharge site flooded above land-surface. The problem was traced to the calibration of the first numerical layer. For the model to produce a good fit to the observation data in the second numerical layer, PEST had adjusted the vertical hydraulic conductivity of the first layer far below what could reasonably be expected. Several further calibration runs were undertaken; however, all reproduced the same problem. Recalibration, including the trial data, resolved this problem. Simulating MAR recharge during calibration relied on the assumption the trial had reached equilibrium within the first few months of operation immediately surrounding the recharge basin. The assumption was supported by analytical modelling. The model was successfully recalibrated including MAR recharge without causing the cells to flood above topography. Overall, a satisfactory fit was achieved to the observation data, with the model error approximately 4% (Table 5-5, Figure 5-13), and the total mass balance was in line with the conceptual understanding (Table 5-6).

Table 5-5 Pilot point calibration statistics

Statistical measure	*Base model (m)*
Residual mean	2.06
Residual standard deviation	9.13
Absolute residual mean	6.37
Residual sum of squares	8.58×10^3
RMSE	9.36
Minimum residual	-31.29
maximum residual	31.29
Range of observations	253.12
Scaled residual standard deviation	0.036
Scaled absolute mean	0.025
Scaled RMSE	0.037
Number of observations	98

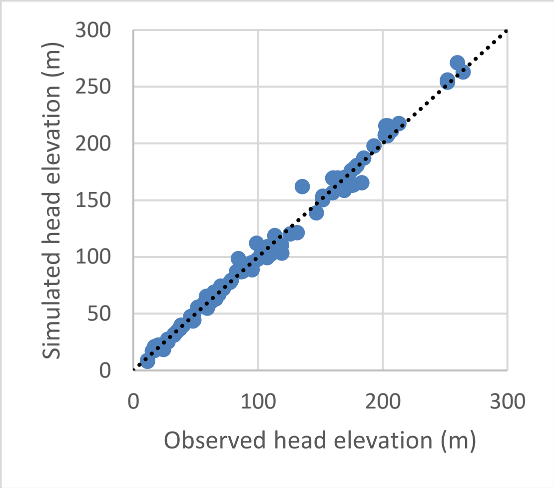

Figure 5-13 Calibration fit.

Table 5-6 Heterogeneous model mass balance

Mass balance	*Inflow m^3/day*	*Outflow m^3/day*
MAR	9000	
Fixed head boundary	11917	380941
General head boundary	3832	
Stream package	205230	26
Recharge	404694	
Drain package		18650
Groundwater abstraction		235057
Total	634674	634675
Percentage error	-4.7x10-5	

Unlike the homogeneous model, the water table remained below the drain elevation in the lower catchment (Figure 5-14). Failure to intercept the drain elevation caused the estimated mounding to propagate further towards the coast. The homogeneous model, however, demonstrated that the coastward propagation of the mound would cease once the drain elevation is reached by the rising water table. Suggesting that the mounding inside the drainage area would only occur under dry climatic conditions, and once the water table rises to its historical elevation, the coastal drains would prevent the mounding from occurring. The water that would have otherwise mounded would be instead removed from the study area as increased spring discharge. As with the homogeneous model, the inland boundary type was varied to investigate its effects on the estimated mounding. Again, using a fixed head boundary, as opposed to a general head boundary, resulted in less propagation of the mound inland (Figure 5-15).

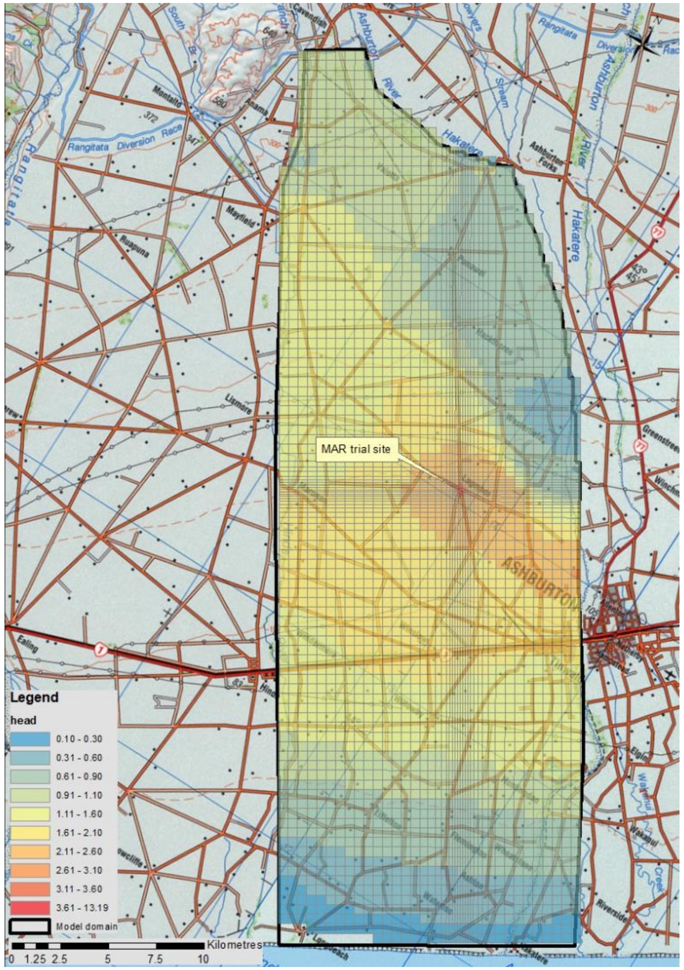

Figure 5-14 Heterogeneous model water table mounding (general head boundary for the inland boundary).

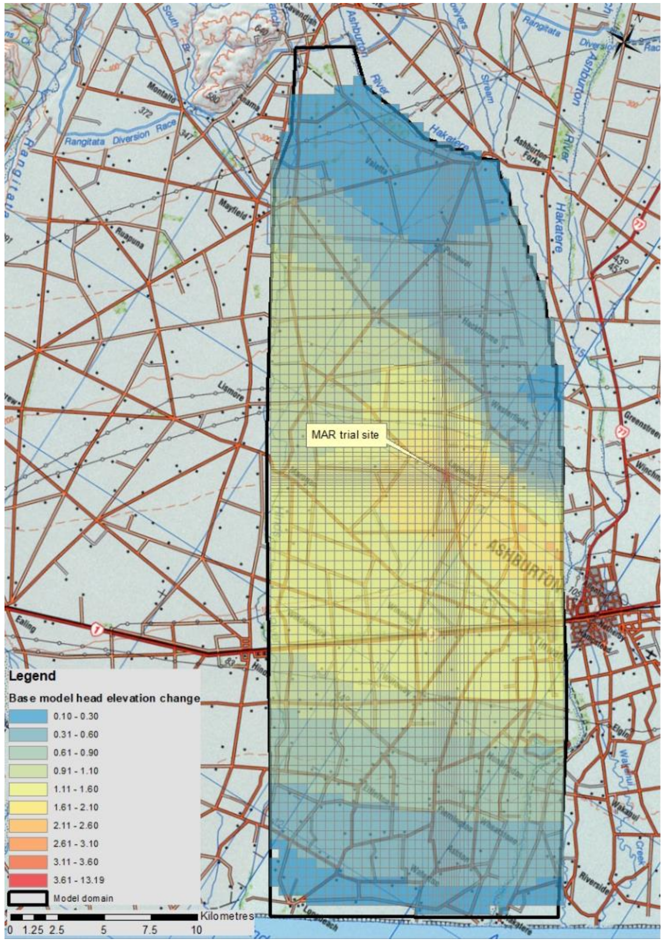

Figure 5-15 Heterogeneous model water table mounding (fixed head boundary for the inland boundary).

Transport modelling was conducted using MT3DMS. Porosity was manually calibrated and like the homogeneous parameter model, indicatively lies in the range of 0.01 to 0.06. Overall, a poor fit to the observed concentration data was achieved. In general, the modelled plume moved too far to the south, missing the two observation wells (BY20/0150 and K37/0200) that showed a clear response to the trial. Carniato et al. (2014) highlighted that transport results are highly dependent on the adopted flow parameters and recommended that calibration of flow models should include both head and transport observations.

5.2.5 Simultaneous calibration of flow and transport

With the findings of Carniato et al. (2014) in mind, a further modelling exercise was conducted, involving recalibration of the highly parameterised model using the observed concentration data to constrain the calibrated hydraulic conductivity fields.

Pilot point PEST was used to automate calibration. Initial parameters were sampled from the homogeneous hydraulic conductivity model. Again, 98 head elevations and 5 flux targets were employed. Twelve concentration observation targets were used. Because only 12 concentration targets were used, they were given high weightings to enforce calibration prioritisation. Calibration used SVD assist to focus the calibration on the highly sensitive parameters. Porosity (homogeneous) was introduced as a calibration parameter and MODFLOW and MT3DMS were run in series. Both the flow model and transport model calibration statistics were combined in the RMSE of the objective function. The inland model boundary was set as a general head boundary for initial calibration and then transformed into a fixed head boundary for subsequent analysis.

Recalibration produced a more satisfactory model. Calibration to groundwater levels and flux discharges was slightly improved, and an improved fit to the concentration data was achieved. The mass balance was still inside the bounds of the conceptual model, though sitting at the lower end (Table 5-8). Mounding estimated from this calibration is shown in Figure 5-16.

Table 5-6 Simultaneously calibrated model mass balance

Mass balance	*Inflow m^3/day*	*Outflow m^3/day*
MAR	9000	
Fixed head boundary	1853	153837
Stream package	126809	1771
Recharge	404452	
Drain package	0	149220
Groundwater abstraction	0	237279
Total	542114	542107
Percentage error	0.001	

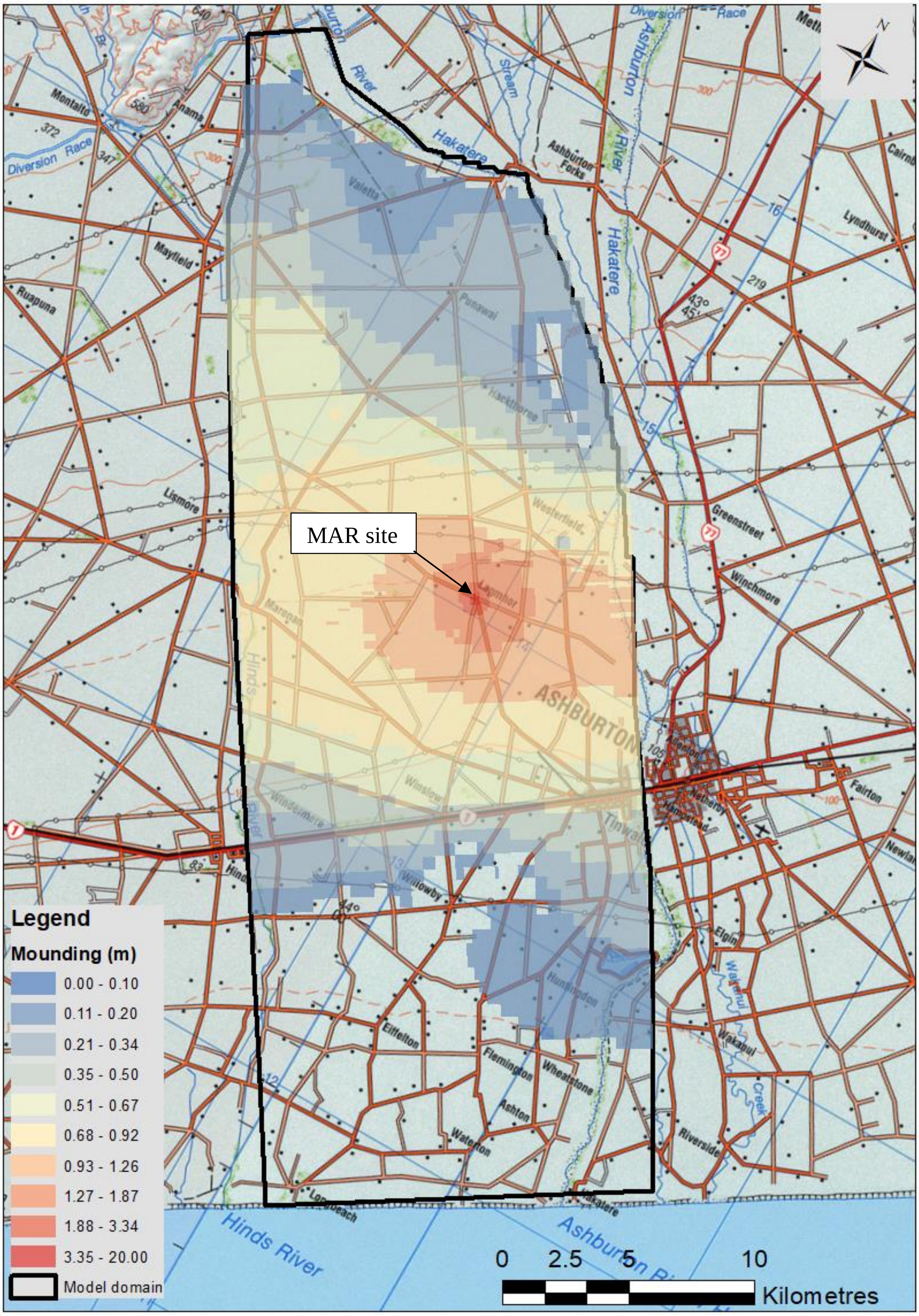

Figure 5-16 Modelled groundwater level change in response to the MAR trial.

The modelled mound was more pronounced than that produced by the head and flux target only model, though the general shape is similar. The MAR trial results in an increased discharge to the drainage network of approximately 7900 m^3/day, nearly 88% of the total MAR trial volume. The transport calibration is comparable to that of the homogeneous model Figures 5-17 through 5-20. The porosity estimated by PEST is approximately 0.04.

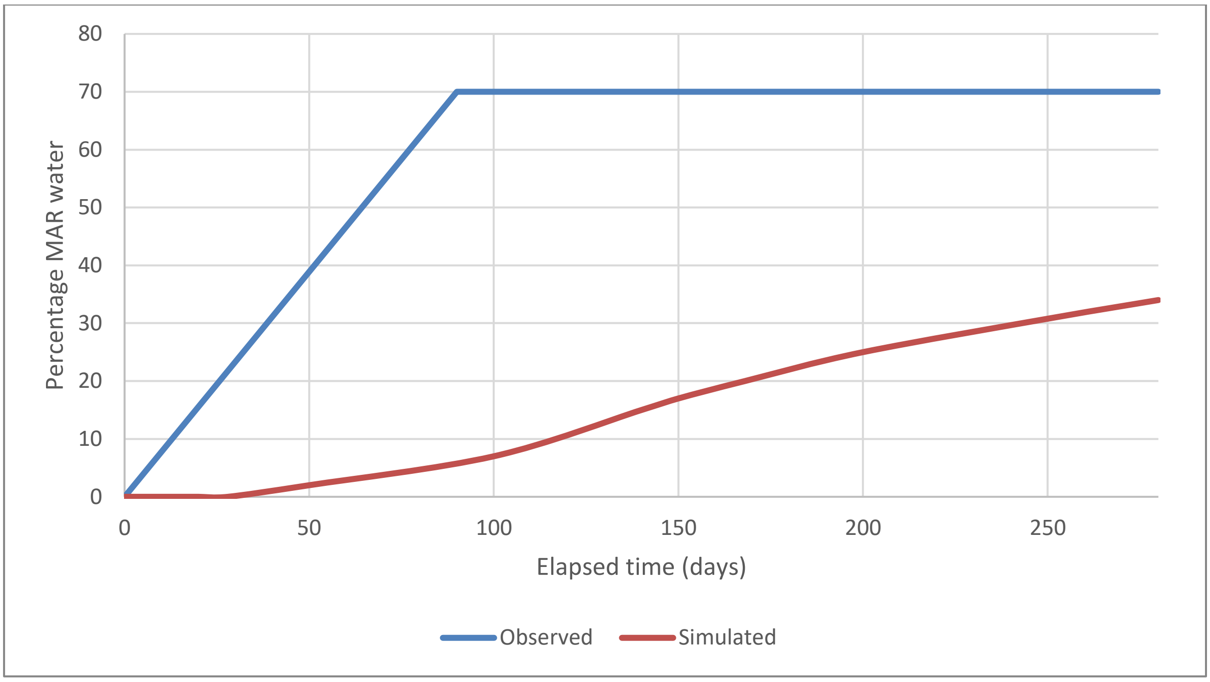

Figure 5-17 Plume breakthrough BY20/0152 (porosity 0.04).

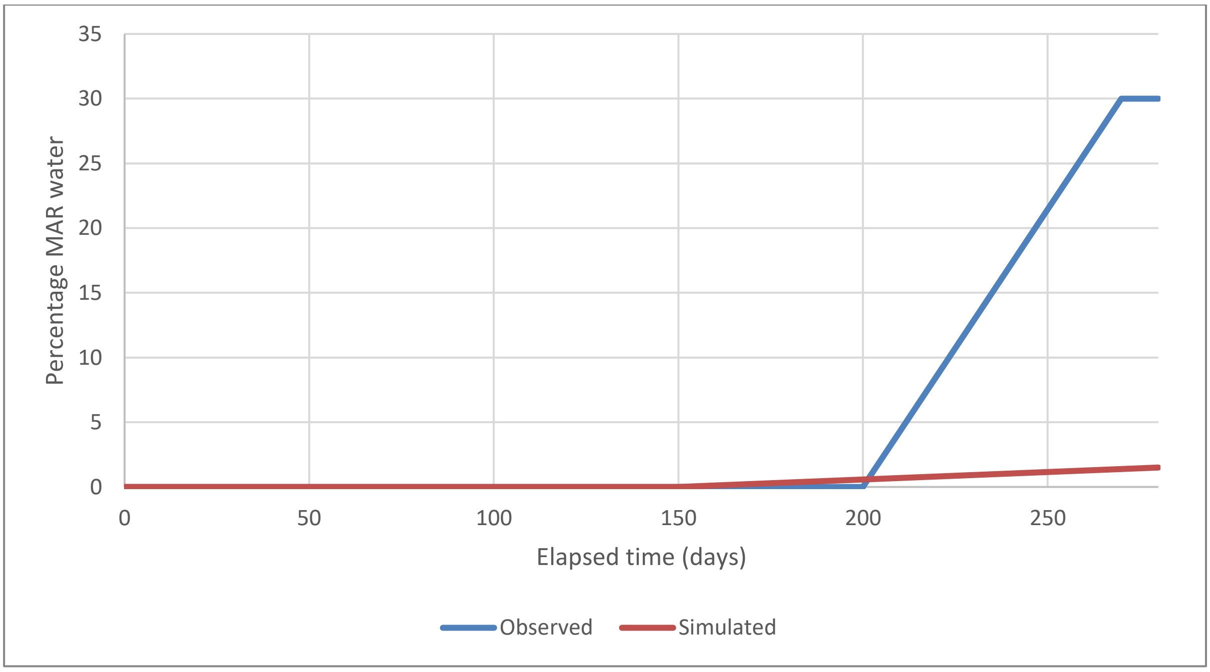

Figure 5-18 Plume breakthrough K37/0200 (porosity 0.04).

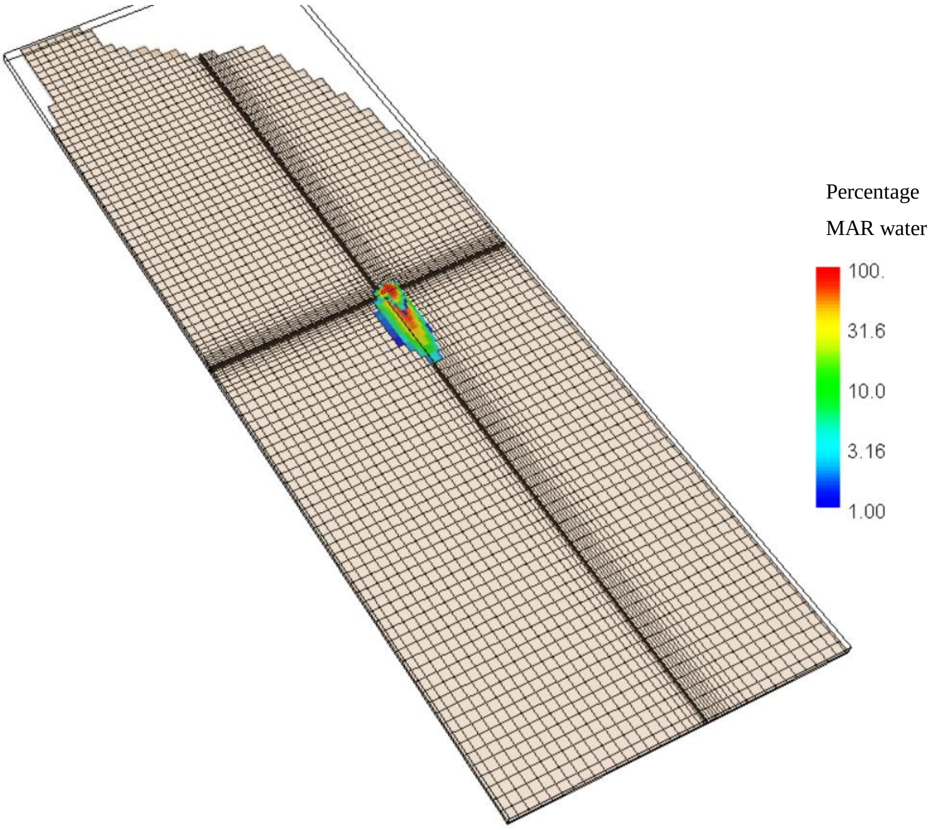

Figure 5-19 Percentage of water from MAR plume after 280 days.

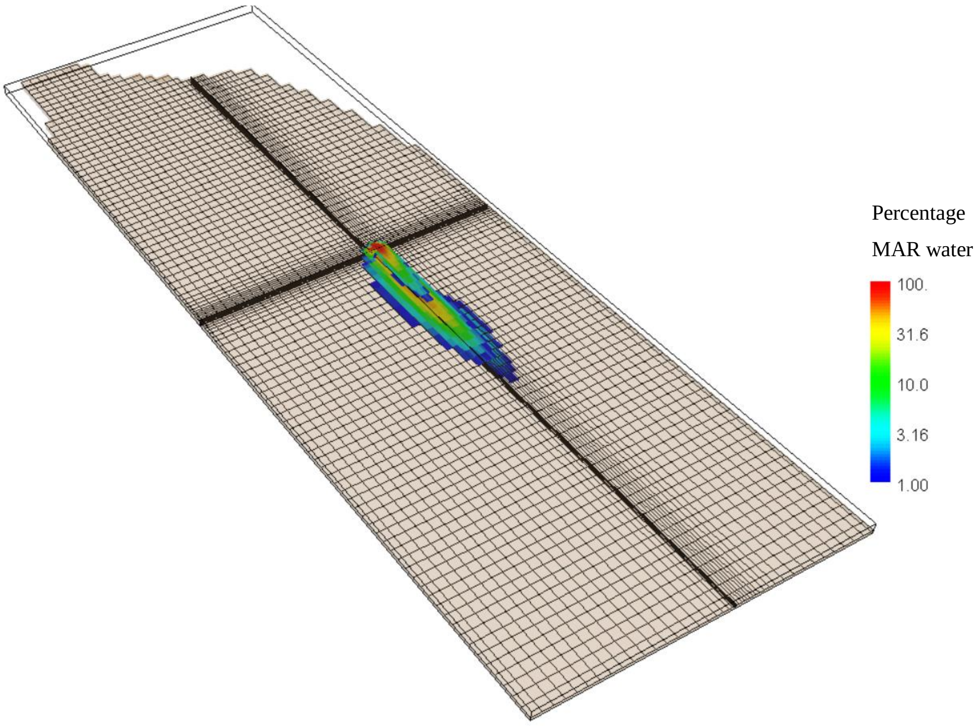

Figure 5-20 Percentage of water from MAR Plume after 1825 days.

5.3 Uncertainty analysis MODFLOW Null Space Monte Carlo

Two rounds of uncertainty analysis were run on the Pilot Point calibrated heterogeneous model. The first round of uncertainty analysis used the NSMC approach on the head elevation and flux calibrated model. Due to computational requirements, only 100 NSMC simulations were produced. The second round used NSMC uncertainty analysis on the head, flux and concentration calibrated model, again a further 100 simulations were run.

With the head and flux calibrated model, while the calibration statistics were approximately equal to those of the base model, the suite of models failed to capture the transport of the MAR plume accurately. The results are presented in Appendix D. The second suite of NSMC results was able to more accurately reflect the water quality changes observed in response to the trial. Run times for the head, flux and concentration calibrated model were 10 times longer than the initial heterogeneous model. Despite this, a further 100 iterations were completed. The parameter bounds were set looser than those of the original simulation at 0.25 and 4 times the calibrated hydraulic conductivities. Only one recalibration cycle was allowed.

The resultant calibration statistics, like the revised base model, were significantly better than with the initial NSMC suite. They represent a significantly broader range of parameters and are, consequently likely to better sample the true population of possible parameters. Table 5-9 shows the calibration statistics.

Table 5-7 Calibration statistics

Statistical measure	*Base model (m)*	*5th percentile (m)*	*50th percentile (median) (m)*	*95th percentile (m)*
Residual mean	-0.15	-0.69	-0.29	3.88
Residual standard deviation	6.13	5.98	6.05	6.87
Absolute residual mean	4.06	4.07	4.29	5.55
Residual sum of squares	$3.69x10^3$	$3.55\ x10^3$	$3.59\ x10^3$	$6.1\ x10^3$
RMSE	6.13	6.02	6.05	7.89
Minimum residual	-26.31	-24.31	-22.27	-14.23
Maximum residual	18.04	15.37	18.41	27.61
Range of observations	253.12	253.12	253.12	253.12
Scaled residual standard deviation	0.024	0.024	.024	0.027
Scaled absolute mean	0.016	0.016	0.017	0.022
Scaled RMSE	0.024	0.024	0.024	0.031
Number of observations	98	98	98	98

Statistics for mounding on a cell by cell basis were generated to enable assessment of the possible mounding at any given location. The extent of mounding can vary considerably from the base model. The median results and average results are very similar to the base calibration. However, the propagation of the mound under the 5th percentile is several kilometres less, while the 95th percentile mounding extends several kilometres further (Figure 5-21) when considered on a cell by cell basis. The 5th and 95th percentiles of mounding suggest the total area with mounding greater than 0.1 m may range from 378 km^2 to 548 km^2.

In terms of MAR effects on the mass balance of the model, the NSMC suite suggests that the drainage network flows will increase by 82% to 88% of the MAR recharge rate. The rest of the MAR recharge is split between offshore discharge and reduced recharge from the Ashburton and Hinds rivers.

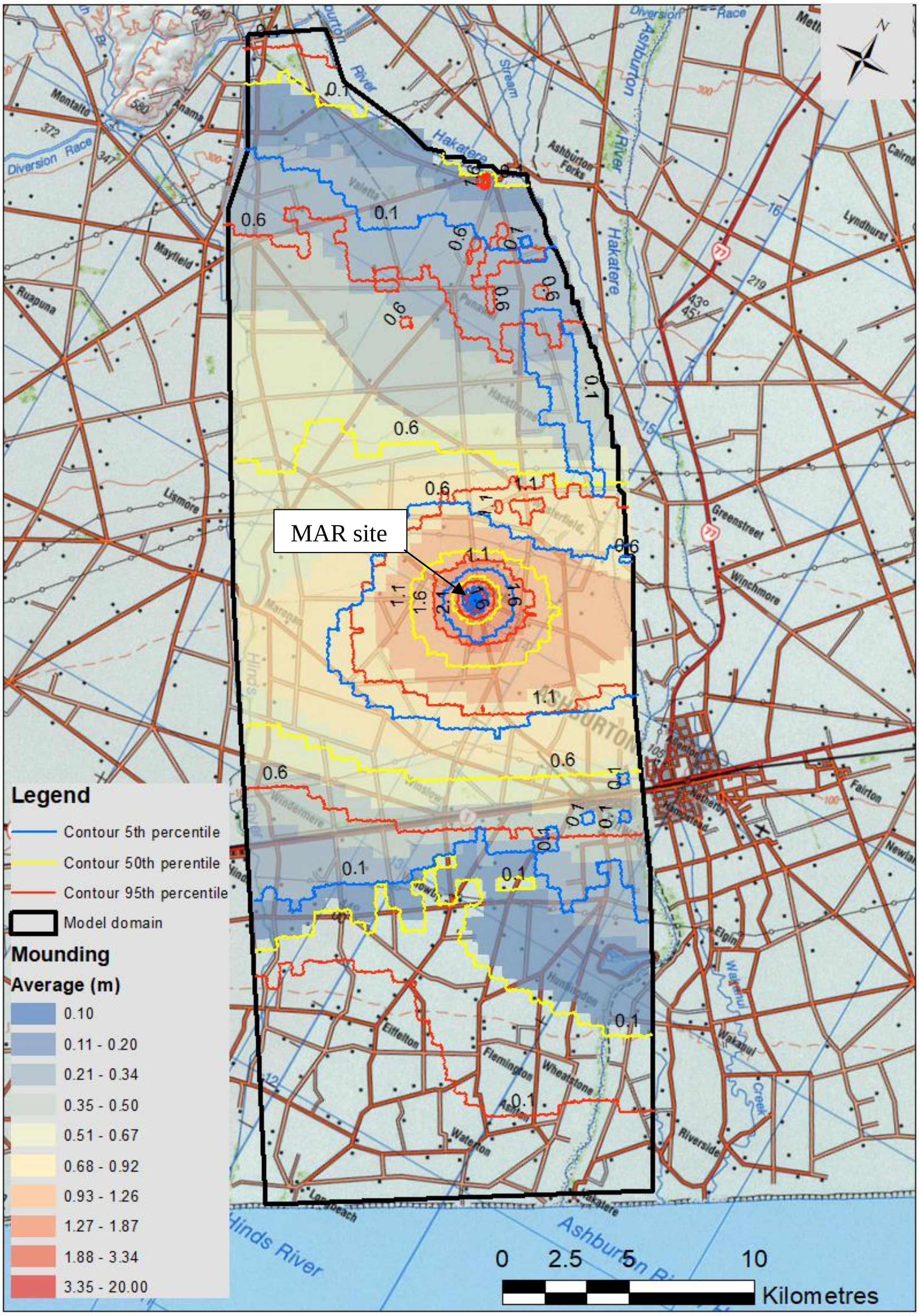

Figure 5-21 NSMC estimate of the extent of mounding possible at any given location. Representing the mounding possible at any given cell, rather than the overall mounding estimated by a particular model run.

5.3.1 Transport model results using 5th, 50th and 95th percentile flow fields

Water quality transport was added to the NSMC generated flow models using the USGS software MT3DMS. While the transport models used a constant flow field; they were themselves run as transient simulations. Three flow fields representing the 5th, 50th and 95th percentile mounding were used in the transport modelling. Each transport model was run across five years, with results reported at the end of the five-year period.

Porosity was calibrated simultaneously with hydraulic conductivities. Automated calibration with PEST estimated the porosity between 0.02 and 0.07. Figures 5-22 to 5-28 show the propagation of the averaged MAR plume using the 5th, 50th and 95th percentile flow fields. PEST calibrated porosities were used for each flow field.

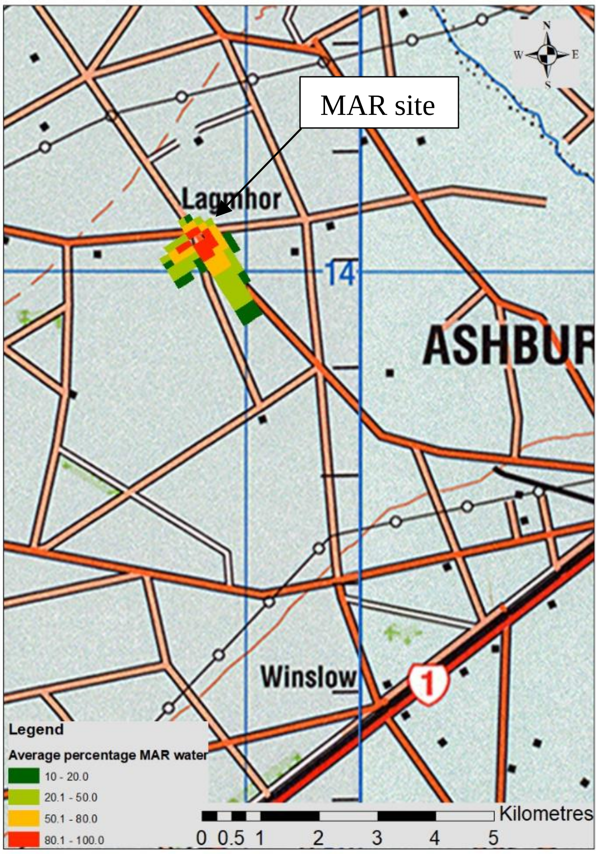

Figure 5-22 MAR plume Layer 2.

Figure 5-23 MAR plume Layer 3.

Figure 5-24 MAR plume Layer 4.

Figure 5-25 MAR plume Layer 5.

Figure 5-26 MAR plume Layer 6.

Figure 5-27 MAR plume Layer 7.

Figure 5-28 MAR plume Layer 8.

6 Discussion and Conclusion

6.1 Analysis of results

Analysis of the study results has been split according to the methods applied. The results have been synthesised to assess whether the trial is likely to be successful in meeting its stated goals of raising groundwater levels, improving spring-fed stream flows and improving both surface and groundwater quality. Finally, the applicability of the modelling approaches is then discussed.

6.1.1 Analytical modelling

In general, the stochastic analytical modelling produced results that were too broad to provide useful results beyond the immediate study area. The stochastic ensemble of mounding suggested that overall, the trial should have limited effects on the surrounding environment. However, when the results are constrained by the analytical assessment of aquifer parameters, the overall zone of influence was significantly larger. The constrained assessment suggested results more like those later produced by numerical modelling. Overall, compared to the stochastic ensemble, the constrained parameters produced a mounding response closer to the 95th percentile than the mean.

6.1.2 Numerical modelling

Both the homogeneous and heterogeneous calibrated models produced more reasonable results than the analytical models (including the constrained results) when considered across the model domain. This is because the overall hydrogeology was too complex across a wide area for the analytical models, and more observation data had been utilised in the numerical model calibration. Both models indicate that the mounding would extend significantly further than anticipated by analytical modelling. However, the extent of the mounding near the trial site was within the range suggested by the analytical modelling.

Homogeneous hydraulic conductivities

The homogeneous hydraulic conductivity model achieved a reasonable fit to the observation data in terms of head elevations and stream discharges. Discharge into the drainage network in the lower catchment was well-calibrated. However, head elevations, particularly in the first numerical layer, were significantly under-predicted in the inland extent of the model domain and over-predicted near the trial site. Adding transport constraints to the calibration scheme made it possible to calibrate the transport model to the observations, without altering the hydraulic conductivity field. Equally, good calibration to groundwater level elevations and flux could be achieved with different combinations of parameters. However, when these hydraulic properties were applied to transport modelling, not all combinations produced a satisfactory result for concentration changes. For instance, an equally well-calibrated result to head and flux was produced using higher vertical conductivity. However, when transport was added to these models, the water quality results were not possible to be calibrated to observations through alteration of porosity alone. The sensitivity of transport results to the modelled hydraulic conductivity fields highlights the need to include water quality observations during flow calibration if transport is later being considered.

Heterogeneous hydraulic conductivities (calibrated to flux and groundwater levels only)

Pilot point calibration of the heterogeneous hydraulic conductivity produced a better fitting model and produced a smaller overall statistical model error, specifically regarding groundwater level elevations when compared to the homogeneous model. This is not an unexpected outcome, considering the high parameterisation of the heterogeneous model.

The initial heterogeneous model performed worse against the estimated spring discharges in the lower catchment, a result of greater weighting having been given to groundwater level elevations. Lower weighting was given to the spring discharges, as the calibration targets were estimates rather than actual known values. Because the spring-fed discharges were a qualitative estimate, it is not possible to determine which calibration scheme was performing better quantitatively in the lower catchment of the model domain. However, qualitatively, it is likely that the homogeneous model was performing better in this region of the model, as it fits better with the conceptual model. For this reason, when the heterogeneous model was recalibrated to include water quality observations, greater weighting was given to the drain discharge targets.

When transport was modelled, the heterogeneity introduced by the pilot points lead to the transport of the MAR plume deeper into the groundwater system, than observations suggest. Overall, there are too few deep observations to show explicitly where the MAR plume will go. However, there are enough observations to show where it will not go. This outcome was likewise encountered with the homogeneous parameters when the sensitivity to hydraulic conductivity was investigated and appeared to be due to unreasonable vertical hydraulic conductivities.

Null Space Monte Carlo uncertainty testing of the head and flux calibrated model

NSMC was employed as a mechanism to address the uncertainty in the modelled groundwater mounding results, introduced by hydraulic conductivity uncertainty. Two rounds of uncertainty analysis were conducted, each generating a total of 100 simulations. The first round of simulations was run through two recalibration iterations before model termination. Calibration statistics from the NSMC runs indicate that each simulation was equally as well-calibrated to observation data, as the initial base model. Mounding assessment used the 5th, 50th and 95th percentiles and average mounding to assess the range of possible outcomes of the MAR trial. Overall, these percentiles produced roughly the same mounding response with generally only a few hundred metres difference (usually only one model grid cell) in mound propagation between the 5th and 95th percentile. The minor differences were somewhat surprising and suggested that the model results were too constrained. Generally, only one recalibration iteration is utilised before model termination for each uncertainty run. However, in this instance two iterations were allowed, as the initial iteration failed to return the modelled calibration; the result, however, appears to be that the true uncertainty in hydraulic conductivity has not been fully addressed.

The water quality outcomes were as unsatisfactory as those using deterministic heterogeneous parameters and highlight that the true aquifer responses were not captured even with the NSMC suite. This means greater uncertainties associated with the hydraulic properties of the aquifer exist than are estimated by the NSMC suite. As with the deterministic approach, the transport scenarios show the MAR plume infiltrating rapidly to depths beneath the screen intervals of wells that showed influence from the trial. Also, like the deterministic approach, the greatest changes in the model are simulated in layer 4, whereas they should have manifest in layer 2 of the model. Further, the water quality outcomes from the NSMC calibrated models suggest that the plume of MAR water would head slightly to the south-southeast, while observation data from well K37/0200 suggests that at least some of the movement is towards the south-east, following the topography. Overall, the NSMC simulations like the deterministic heterogeneous parameter calibrated model appeared to perform worse than the homogeneous parameter model, at replicating the observed changes in water quality in response to the trial.

Failure of the NSMC simulations to adequately capture the transport of the MAR plume suggests that the true range of parameter uncertainty was not sampled. The logical conclusion, therefore, is that the derived percentiles from the NSMC simulations were not, in fact, a true representation of the possible outcomes of the MAR trial. Instead, they offer only a few variations centred around an unknown statistical probability. Previous research has demonstrated similar findings, suggesting NSMC results, especially when applied to transport related problems can be biased because of their basis on a single calibration. Research has demonstrated that bias occurs because NSMC PEST can end up focusing on local calibration optima and fail to capture the global optima (Yoon et al. 2013). Yoon et al. (2013) suggest that an effective mechanism to address this is to utilise multiple calibration constrained NSMC (MNSMC). While MNSPC has not been undertaken in this research due to the deficiencies in the initial uncertainty assessment, further calibration was conducted, as discussed below.

Simultaneous calibration of flow and transport

Recalibration of the highly parameterised model with the inclusion of transport targets was able to produce a significantly improved calibration to head, flux and concentrations targets. Further, the results fit better with the initial conceptual model. The improved performance of the flow model constrained by transport observations is not surprising. Improved calibration by this method is well documented (Carniato 2014). The improved predictive performance relative to the initial NSMC simulations highlights that constrained uncertainty analysis undertaken without the guidance of transport observations was insufficient at capturing the full range of possible hydraulic conductivity combinations.

It is suggested that the data types used to constrain the uncertainty analysis are what the uncertainty analysis itself is useful for assessing. Therefore, to provide meaningful conclusions for transport-related questions, it is suggested that transport observations are used in the calibration of the base model and later uncertainty assessments by NSMC.

Null Space Monte Carlo uncertainty testing of groundwater level, flux and transport calibrated model

The second round of NSMC analysis investigated the uncertainties in trial results, using the flow and transport calibrated model. Unlike the initial NSMC analysis, the mounding results of the simulations show there is greater variance in the predicted mounding. Qualitatively, this provides more confidence that the results in terms of mounding are well represented. Unlike the initial NSMC simulation, using concentration targets to constrain the calibration has enabled the NSMC runs to produce transport observations that better match the first two years of trial results. Mounding assessment used the 5th, 50^{th} and 95^{th} percentiles and average mounding to assess the range of possible outcomes of the MAR trial. Overall, statistical percentiles show that the extent of mounding greater than 0.1 m may vary by as much as five kilometres, but will, in general, be truncated down-gradient by the drainage network, which removes between 82% and 88% of the recharge volume. Transport modelling associated with the NSMC suite, while improved compared with the initial assessment, still suggests the MAR plume would travel slightly further to the south than was indicated by observation data. Despite this, the results can sufficiently answer whether MAR will positively impact groundwater and surface water quality.

Comparison between the results from homogeneous and homogeneous hydraulic conductivities

When considering questions related to mounding and transport at a regional scale, homogeneous parameters may be enough to show the distribution of groundwater level changes and water quality because of MAR. Homogeneous modelling is significantly faster to calibrate and produce results comparable with heterogeneous modelling, and the uncertainties associated with it can be outweighed by the easy application of the approach. However, when it comes to site-specific examples, the homogeneous model performed worse in terms of calibration and history matching than the heterogeneous model. Both approaches would likely lead to erroneous conclusions around water quality outcomes if transport was not calibrated concurrently with flow.

6.2 Effectiveness of the Hinds MAR trial at reaching stated goals

6.2.1 Raised groundwater levels

Modelling suggests that groundwater levels over a large area are likely to be raised in response to the trial. However, a review of the fluctuations in water levels associated with rainfall (Section 3), shows that it would be difficult to directly identify the effect of MAR at a distance of more than several kilometres down-gradient from the trial. Modelling suggests mounding immediately around the trial site ranges from 5 m to 15 m. This compares to the seasonal variations of 10 m observed in control well K37/0215 and 25 m observed in monitoring well BY20/0151. Further down-gradient, the NSMC suite of simulations suggest that mounding greater than 1 m may propagate between 2.5 km (5^{th} percentile) and 5 km (95^{th} percentile) from the recharge site. Likewise, mounding greater than 0.5 m may propagate between 5 km and 9 km down-gradient under the 5^{th} and 95^{th} percentiles respectively. Finally, mounding greater than 0.1 m might extend between 10 km and 20 km below the recharge site. These mounding responses would be very difficult to discern amongst the seasonal water level variations that can exceed 20 m.

6.2.2 Raised flow in the lowland streams

Analytical modelling is unable to predict whether the changes in groundwater mounding will have a positive impact on spring-fed stream discharge in the lower Hinds plains. Results of the homogeneous parameter calibrated modelling suggested the trial will be successful in increasing spring-fed stream discharges in the lower catchment of the study area. The heterogeneous modelling likewise suggest that the MAR trial will increase discharge to the drainage network. Increases in discharge may account for as much as 88% of the MAR recharge volume.

6.2.3 Water quality improvements

All numerical modelling scenarios suggest the plume of MAR water will propagate several kilometres coastwards of the trial site. Results suggest that the plume may extend to near SH 1 after 5 years of operation. Overall, the findings suggest the MAR trial will successfully dilute the aquifer immediately down-gradient from the trial site. However, there is a discrepancy in the results as to whether the plume will move to the south, to the southeast or more due southeast. Both the base model, the 95^{th} percentile flow field and median flow field suggest the MAR plume will travel deeper into the aquifer than observation data suggests. However, the 5^{th} percentile flow field and the manually calibrated results from the homogeneous parameter model, suggest the plume will stay shallower in the groundwater system. There is an agreement between all the models that the plume will not reach the spring-fed streams during the duration of the trial. While the model suggests the MAR plume would propagate too deep into the aquifer to cause improvements in the surface water quality, there is still high NO_3-N contamination at depth, which might be diluted by the MAR trial. NO_3-N is encountered at concentrations of greater than 6 mg/l at depths of over 100 m below ground surface in the lower plains.

Overall, the modelling suggests that inside the plume extent, MAR may have a significant effect in terms of dilution. Inside the plume, concentrations up to 3 km down-gradient from the recharge basin may be diluted as much as 80%. This means where the initial concentration is 15 mg/l NO_3-N, the final concentration once the MAR plume breaks through can be expected to be approximately 3 mg/l NO_3-N. The study results suggest that 50% dilution may extend as far as 8 km down-gradient, while 20% dilution may occur as far as 11 km from the trial site after 5 years.

6.3 Applicability of each modelling approach

6.3.1 Analytical mounding models

Analysis of the analytical modelling suggests that it can be useful at the local scale. When scaled up to a larger area, such as the model domain, the lack of constraint to known aquifer parameters or head observations across the catchment results in mound propagation that appears unreasonable. Analytical models, like numerical models, suffer from what is known as equifinality, meaning many combinations of parameters can produce results that fit the observed data equally well. This results in significant uncertainty as to the true propagation of the groundwater mounding, as is evidenced by the study's results, which suggest mounding may be limited to within five kilometres of the trial or extend more than 10 km down-gradient. Uncertainty analysis can help constrain the probable outcomes, as can constraining the parameter ranges based on observation. However, over large areas such as this study, the analytical approach is less suitable than numerical modelling.

Further, the limitations of the analytical equations, such as the Hantush mounding equation, mean that analysis by analytical tools is generally only applicable to groundwater mounding and cannot be applied to estimating changes to stream flows because of the recharge water. This limitation extends to water quality modelling as well, with the Hantush equation only capable of assessing groundwater mounds.

Despite the limitations of analytical models, they enable fast and simple analysis of mounding and have very low computational requirements compared to numerical models. Analytical models are suitable for assessing the local-scale effects from recharge trials when the results can be constrained by observation data.

6.3.2 Homogeneous parameter numerical models

The homogeneous parameter calibrated numerical model produced results that are reasonable both in terms of conceptual model and model calibration. Given the computational ease with which homogeneous parameter calibrated models can be produced, they seem to offer a reasonable compromise between computational requirements and reliable results. Further, the ability to reproduce the water quality changes as a result of the MAR trial without the need for recalibration of the hydraulic conductivity fields, suggests there may be a pragmatic approach to assessing MAR programmes in the future. While the uncertainty of homogeneous parameter calibrated models has not been assessed in the study, given their low computational requirements, assessment of uncertainty could take place using Monte Carlo analysis and the GLUE approach. Such an approach would lead to a robust assessment of uncertainties and provide confidence in model results. Overall, homogeneous parameter calibrated models seem to be a good fit for assessing the impact of MAR at the regional scale. However, homogeneous parameter models are less efficient; compared with heterogeneous parameter calibrated models, at reproducing the exact head elevations and changes at specific points in the model domain.

6.3.3 Heterogeneous parameter numerical models

The highly parameterised pilot point PEST calibrated numerical model produced the best statistical and spatial fit to groundwater level across the model domain and performed as well qualitatively against the conceptual model. Pilot point PEST calibrated models are significantly more computationally demanding to develop than homogeneous parameter models, and the value of added complexity must be commensurate with the computational costs.

The failure of the heterogeneous model to perform as well as the homogeneous model, in terms of transport model fit, without recalibration including concentration targets, suggests the extra effort of pilot point calibration may not be justified if the questions being asked of the model are at a catchment scale. However, the computational effort does seem justified for assessing local-scale responses to MAR. The failure to estimate the actual spatial distribution of water quality changes highlights that calibration to groundwater levels and flux only is not enough to capture the true spatial distribution of hydraulic conductivities. If heterogeneous modelling is to be used to aid in answering transport-related questions, then water quality observations need to be included in the calibration criteria and transport modelling calibrated concurrently.

6.3.4 Null Space Monte Carlo

Pilot point PEST and NSMC present a stochastically robust way to quantify the uncertainty in model results. Being computationally less demanding than a highly parameterised brute force Monte Carlo approach with utilisation of GLUE, it offers many advantages to decision-makers over straight deterministic modelling. However, as has been demonstrated in this research, it is not enough to remove the need for the utilisation of water quality data as a calibration constraint when considering transport-related problems.

In this research, NSMC has successfully been employed to investigate the range of calibration constrained predictions of groundwater mounding responses to the Hinds MAR trial. However, when these models were applied to water quality transport modelling, in the absence of the original model being calibrated to water quality constraints, they failed to produce the pattern of water quality responses observed in 3D. Subsequent recalibration, including concentration targets, produced more satisfactory results and highlighted the usefulness of NSMC.

6.4 Recommendations

- For local scale mounding studies, analytical models provide a simple and robust mechanism for assessing possible effects. However, without aquifer test data to constrain the aquifer parameters, the range of possible outcomes is likely too large to be practical beyond a first pass assessment.
- Calibration constrained homogeneous parameter models proved a useful intermediary between analytical and highly parameterised numerical models, as they have demonstrated to be effective at estimating the likely mounding and water quality changes to MAR at a catchment scale.
- Highly parameterised heterogeneous parameter calibrated models produce the best statistical fit to observation data. However, unless the flow component is calibrated simultaneously with transport, they should be limited to assessing quantity changes.
- NSMC presents a robust and relatively straightforward method for implementing uncertainty analysis, however conducting uncertainty analysis on hydraulic conductivities not constrained by transport observations, risks not capturing the range of possible water quality outcomes or being able to match water quality observations.
- If transport-related questions are expected to be answered by the model, then water quality targets should be included to constrain the flow model calibration, as calibration of porosity has been demonstrated insufficient to explain observations.

6.5 Conclusion

The impacts of the Hinds MAR trial have been assessed through modelling, using a variety of approaches ranging from simple analytical models to more complicated numerical models and uncertainty techniques. While all these models have been able to reproduce the effects of the trial in the immediate vicinity, only numerical modelling has been able to reproduce approximations of the regional scale mounding extent.

Numerical models calibrated using homogeneous parameters can reproduce the general scale of effects observed from the MAR trial. However, they fail to capture the head elevation distributions as effectively as the heterogeneous parameter calibrated models. Despite this, when it came to water transport modelling, the homogeneous model performed as well as the heterogeneous models at reproducing the observed spatial distribution of water quality changes. The initial failure of the heterogeneous parameter calibrated models to reproduce the water quality changes were linked to not calibrating the flow model to concentration targets. When the flow model was recalibrated, including concentration data, the model performance was comparable to the homogeneous model. Had more observations been available, the calibration to water quality would likely have improved and may have facilitated calibration of dispersivities.

Like the heterogeneous base model, without including concentration targets, none of the initial NSMC simulations was able to produce the water quality changes observed. In all instances, the plume of MAR water was either transferred too deep in the model structure or moved too far to the south. The second suite of NSMC simulations was better able to reproduce observations.

The need to calibrate flow models to transport observations was perhaps the most significant finding of the study, and like Carniato (2014), showed that development of a transport model post-calibration of a flow model, was unlikely to be successful in reproducing actual observations. It is suggested that key to the adequate capture and representation of MAR trial results is simultaneous calibration of aquifer properties to both head elevations and water quality observations.

Despite the limitations and an inability to reproduce the exact horizontal and vertical distribution of water quality changes, the deterministic and NSMC numerical models were able to answer the key questions of the MAR trial. They have shown it will be possible to raise groundwater levels across a large area, increase stream flows and improve water quality in groundwater. However, the MAR trial will be unlikely to influence surface water quality. All the numerical models created as part of the study show that the trial will raise groundwater levels across a significant area. Further, modelling results suggest that the trial will result in increased spring-fed discharges across the lower catchment. Transport modelling utilising all the flow models suggests improvements in water quality will be seen for several kilometres down-gradient of the trial site. The transport modelling suggest improvements in the water quality of the coastal drains are unlikely, however, if very low porosity exists, improvements are possible. Only those simulations utilising the lower bounds of reasonable porosity, would propagate that distance to the drains within the five years of the trial operation. However, the lower bound porosities do not seem to be representative of the aquifer at the catchment scale.

In terms of the appropriateness of the various modelling approaches for assessing MAR trial outcomes, the study shows that analytical models, while fast to build and useful at the local scale, are less appropriate when considering mounding at the catchment scale in fast-flowing alluvial aquifers, due to simplistic assumptions and sensitivity to the adopted conceptual model. Further, analytical models, such as the Hantush mounding equation cannot be used to assist water quality changes or spring discharge changes, meaning separate analytical models would be required to answer these questions.

The homogeneous parameter calibrated numerical model was able to reproduce a reasonable fit to head elevations, and spring discharges across the model domain conserving the conceptual water balance. It was also successful at reproducing the spatial distribution in both the vertical and horizontal axis for the transport model. At a catchment scale, the homogeneous parameter calibrated modelling approach seemed to be sufficient to enable understanding of the outcomes of recharge trials.

Pilot point calibration of heterogeneous hydraulic conductivity parameters was able to reproduce a better spatially distributed calibration and a better overall calibration statistic against the observed groundwater level elevations and flux, when compared to the homogeneous model. So long as water quality data are used in the calibration of the flow model, a reasonable approximation of the observations could be produced. However, this was not the case when concentration targets were not initially used in the flow model calibration. Use of NSMC has provided confidence in the probable outcome, but again, the base model must be calibrated to concentration data if water quality questions are asked of the study.

In summary, for further MAR trials, analytical modelling is likely sufficient to estimate the mounding effects within the immediate vicinity of the trial. However, if catchment scale questions are to be answered, numerical modelling should be applied. Where numerical modelling is to consider transport problems, then water quality observations should be included in the calibration targets, and flow and transport calibrated simultaneously. Calibration to flow and uncertainty analysis alone was not sufficient to capture the possible spatial distribution of water quality changes. However, uncertainty analysis was useful for providing confidence in modelled outcomes and should be employed if and when the risk associated with MAR programmes needs to be considered.

7 References

Barrell D, Forsyth P, McSaveney M (1996) Quaternary geology of the Rangitata Fan, Canterbury Plains, New Zealand. Lower Hutt

Beven K, Binley A (1992) The future of distributed models: Model calibration and uncertainty prediction. Hydrol Process 6:279–298. doi: 10.1002/hyp.3360060305

Beven K, Binley A (2014) GLUE: 20 years on. Hydrol Process 28:5897–5918. doi: 10.1002/hyp.10082

Bidwell V (2005) Realistic forecasting of groundwater level, based on the eigenstructure of aquifer dynamics. Math Comput Simul 69:12–20. doi: 10.1016/j.matcom.2005.02.023

Boyle T (2012) Ashburton River (Hakatere) Flood Hazard Management Strategy

Brown L, Wilson D, Moar N, Mildenhall D (1988) New Zealand Journal of Geology and Geophysics Stratigraphy of the late Quaternary deposits of the northern Canterbury Plains, New Zealand. New Zeal J of Geology Geophys 31:305–335. doi: 10.1080/00288306.1988.10417779

California Water Board (2014) DPH-14-003E GW Replenishment Using RW May 30, 2014. 1–77

Carleton G (2010) Simulation of Groundwater Mounding Beneath Hypothetical Stormwater Infiltration Basins. Sci Investig Rep 2010–5102:1–76

Carniato L, Schoups G, van de Giesen N, Seuntjens P, Bastiaens L, Sapion H (2014). Highly parameterized inversion of groundwater reactive transport for a complex field site. *Journal of Contaminant Hydrology* 173 (2015) 38-58

Cooper H, Jacob C (1946) A generalized graphical method for evaluating formation constants and summarizing well-field history. Trans Am Geophys Union 27:526. doi: 10.1029/TR027i004p00526

Cozby P, Bates S (2012) Methods in behavioural research. McGraw-Hill

Dann L, Close M, Pang L, Flintoff M, Hector R (2008) Complementary use of tracer and pumping tests to characterize a heterogeneous channelized aquifer system in New Zealand

Davey G (2006a) The Effects of Border Dyke Irrigation Recharge on Groundwater Levels in and below the Valetta Scheme Report U06/11

Davey G (2006b) Definition of the Canterbury Plains Aquifers.

Dench W (2017) Identifying Changes in Groundwater Quantity and Quality Resulting from Border-dyke to Spray Irrigation Conversion. A thesis submitted for a degree Masters of Science at Canterbury University

DHI (2017) MIKE SHE Vol 2. Reference Guide

Dillon P, Kumar A, Kookana R, (2009) Managed Aquifer Recharge - Risks to Groundwater Dependent Ecosystems - A Review. Water

Dillon P, Stuyfzand P, Grischek T, (2018) Sixty years of global progress in managed aquifer recharge. Hydrogeol J 27:1–30. doi: 10.1007/s10040-018-1841-z

Doherty J (2003) Ground Water Model Calibration Using Pilot Points and Regularization. Gr. Water 41:170–177

Doherty J (2019) PEST Model-Independent Parameter Estimation - User Manual Parts I and II. Watermark Numerical Computing

Doherty J, Fienen M, Hunt R (2010a) Approaches to Highly Parameterized Inversion : Pilot-Point Theory, Guidelines, and Research Directions. USGS Sci Investig Rep 2010–5168 36. doi: 2010-5168

Doherty J, Hunt R (2010) Approaches to highly parameterized inversion—A guide to using PEST for groundwater-model calibration: U.S. Geological Survey Scientific Investigations Report 2010–5169

Durney P, Ritson J (2014) Water Resources of the Hinds/Hekeao Catchment: Modelling Scenarios for Load Setting Planning Process. Environment Canterbury Regional Council

Durney P, Ritson J, Druzynski A, Alkhaier F, Tutulic D, Sharma M (2014) Integrated catchment modelling of the Hinds Plains : model development and scenario testing. Environment Canterbury Regional Council

Durney P (2019) Ashburton River flow and allocation modelling. Environment Canterbury technical memorandum

Enemark T, Peeters L, Mallants D, Batelaan O (2018) Hydrogeological conceptual model building and testing: A review. J Hydrol. doi: 10.1016/j.jhydrol.2018.12.007

Environment Canterbury (2017) Canterbury Land and Water Regional Plan. Environment Canterbury Regional Council

Freeze R, Cherry J (1979) Groundwater. Prentice-Hall, Inc., Englewood Cliffs

Ghasemizadeh R, Yu X, Butscher C (2015) Equivalent Porous Media (EPM) Simulation of Groundwater Hydraulics and Contaminant Transport in Karst Aquifers. PLoS One 10:e0138954. doi: 10.1371/journal.pone.0138954

Golder (2014) Hinds/Hekeao Plains Subregional Catchment: Managed Aquifer Recharge (MAR) As a

Tool for Managing Water Quality and Quantity Issues. New Zealand Government - Environment Canterbury Regional Council

Golder (2015) Resource consent application and assessment of effects on the environment. Managed aquifer recharge – Hinds Plains catchment. A report produced for Canterbury Regional Council by Golder Associates (NZ) Limited. Golder report 1478110257-004

Golder (2017) MAR Hinds Managed Aquifer Recharge Pilot Trial. Phase 1 Report. Golder Associates (NZ) Limited. Golder report 1538632-7410-R-Rev2

Hanson C, Abraham P (2013) Cross sections of groundwater chemistry through the Ashburton-Rangitata plain. Environment Canterbury Regional Council

Hantush M (1967) Growth and decay of groundwater-mounds in response to uniform percolation. Water Resources Research, v.3, p.227-234

IGRAC (2019) Groundwater in a Changing World. IGRAC Strategy 2019-23. Delft, Netherlands.

Kazner C, Wintgens T, Dillons P (2012) Water reclamation technologies for safe managed aquifer recharge. IWA Publishing, Colchester, UK

Kronast M (2016) Developing a SEEP/W™ modelling approach to design managed aquifer recharge sites on the Canterbury Plains, New Zealand

Landcare Research (2015) S-Map online an interactive mapping and query application for New Zealand soils data. Retrieved from *https://smap.landcareresearch.co.nz/*

Lilburne L, Hewitt A, Webb T (2012) Soil and informatics science combine to develop S-Map: A new generation soil information system for New Zealand. Geoderma 170(0): 232-238

Loucks P, Beek E (2005) Water Resource Systems Planning and Management

Metcalf & Eddy Inc., an AECOM Company, Asano T, Burton F (2007) Water Reuse : Issues, Technologies, and Applications: Issues, Technologies, and Applications. Mcgraw-hill

Maliva R (2015) Managed aquifer recharge: State-of-the-art and opportunities. Water Sci Technol Water Supply 15:578–588. doi: 10.2166/ws.2015.009

Meredith A, Croucher R, Lavender R, and Smith Z (2006) Mid-Canterbury coastal streams: assessment of water quality and ecosystem monitoring 2000 – 2005. Environment Canterbury Regional Council

Meredith A (2014) Ecological Assessment for Hinds Catchment Streams and Waterways. Environment Canterbury Regional Council

Mitchell D (1980) History of the Ashburton – Hinds Drainage District. South Canterbury Catchment

Board

Moore C (1994) West Melton Aquifer Recharge Trial June-September 1993

Neuman S, Witherspoon P (1969) Theory of Flow in a Confined Two Aquifer System. Water Resour Res 5:803–816. doi: 10.1029/WR005i004p00803

New Zealand Government (2017) National Policy Statement for Freshwater Management Amendment Order 2017

Niswonger R, Panday S, Motomu I (2011) MODFLOW-NWT, A Newton Formulation for MODFLOW-2005. USGS reports 44

Hall S (2012) Ashburton District Council Water Investigation Project. OPUS international consultants Ltd

Pattle Delamore & Partners (2007) Report on September 2005 Eyre River Aquifer Recharge Trial

Poeter E, and McCray J (2008) Modeling Water Table Mounding to Design Cluster and High-Density Wastewater Soil Absorption Systems. Journal of Hydrologic Engineering - J HYDROL ENG. 13. 10.1061/(ASCE)1084-0699(2008)13:8(702)

Rahman M, Rusteberg B, Gogu R (2012) A new spatial multi-criteria decision support tool for site selection for implementation of managed aquifer recharge. J Environ Manage 99:61–75. doi: 10.1016/j.jenvman.2012.01.003

Ringleb J, Sallwey J, Stefan C (2016) Assessment of managed aquifer recharge through modeling-A review. Water (Switzerland) 8:1–31. doi: 10.3390/w8120579

Schulze-Makuch D (2005) Longitudinal dispersivity data and implications for scaling behavior. Ground Water, 43(3), 443-456

Scott L (2013) Hinds Plains water quality modelling for the limit setting process. Canterbury Regional Council Report

Sinclair Knight Merz (2009) Managed Aquifer Recharge Feasibility Study. Report prepared for ENVIRONMENT CANTERBURY. 2009

Stedinger J, Vogel R, Lee S, Batchelder R (2008) Appraisal of the generalized likelihood uncertainty estimation (GLUE) method. Water Resour Res 44: doi: 10.1029/2008WR006822

Theis C (1935) The relation between the lowering of the Piezometric surface and the rate and duration of discharge of a well using ground-water storage. Trans Am Geophys Union 16:519. doi: 10.1029/TR016i002p00519

Van Housen J (2015) Modelling the temporal and spatial variation of evapotranspiration from irrigated pastures in Canterbury. A thesis submitted for a degree of Doctor of Philosophy at Lincoln University

WGA (2018) Hinds/Hekeao Managed Aquifer Recharge Trial - Year 2 Final Report

Yoon H, Hart D, McKenna S (2013) Parameter estimation and predictive uncertainty in stochastic inverse modelling of groundwater flow: Comparing null-space Monte Carlo and multiple starting point methods. Water Resources Research Volume 49, Issue 1. January 2013 Pages 536-553.

Appendices

Appendix A Land surface recharge estimate

Recharge to groundwater is critical to maintaining reliable supplies of water for both anthropogenic and natural uses. By maintaining groundwater levels, it provides base flows to rivers, discharges to springs and potable and industrial supplies for people and industry. Groundwater recharge comes primarily from two sources: losses from rivers and from rain infiltration through soil profiles, and to a lesser extent from excess irrigation. Determination of the quantities of recharge from both primary sources is uncertain due to data sparsity.

Recharge through Canterbury’s soils has been investigated for several decades, both for allocation purposes and as part of modelling exercises (Scott 2004; Durney et al. 2014). Various approaches have been tried using numerous modelling packages that have ranged from Fortran (Scott 2004) and Python scripts, to dedicated hydrological models, such as Hydrus 1D (Van Housen 2015) and MIKE SHE (Durney et al. 2014). Modelling approaches have ranged from soil moisture balance models to full Richards (1931) equation.

As a starting point, the Durney et al. (2014) recharge model of the Hinds and Ashburton area was reviewed to assess accuracy in both timing and the annual volume of recharge. Groundwater level fluctuations in the area south of the Ashburton River/Hakatere (Hinds) were investigated, as was recharge in the three lysimeters located north of the river. Review of groundwater level fluctuations revealed that the existing recharge model of the Hinds failed to replicate the timing of recharge events. With the model generally demonstrating a delayed response of several months, compared to the observation data. Overall, modelled annual recharge volumes were within the range of measured recharge at the Winchmore Lysimeter site.

The 2014 recharge model, utilised Richard’s (1931) equation for unsaturated flow (Durney et al. 2014). The delayed timing of modelled recharge was traced to the use of Richard’s equation on the light and very light Canterbury soils. Due to the suction and wetting front characteristics in Richard’s equation, the modelled recharge was occurring over a more extended period than observed in the field. Incorrect assumptions due to lack of accurate soil physics data appear to be the primary cause of the delay in modelled recharge. Review of other available recharge models suggested that a simple soil moisture balance model could produce reasonable results using readily available soil property information.

Soil properties

To develop the new two-layer recharge model, the available literature on Canterbury soils were reviewed and lysimeter data sourced for calibration. The primary source data for the soil profiles came from S-MAP (Landcare Research 2014), Lilburne et al. (2012) and Van Housen (2015). Aqualinc Research Ltd provided additional soil and neutron probe data.

Aggregates of the soil profile data from Aqualinc, where in general agreement with S-MAP, were used as a starting point for field capacity and wilting point of the soils at the various dryland lysimeters in the Canterbury region. S-MAP provided the estimates of maximum saturated hydraulic conductivity. Table A-1 shows details of the soils mapped by S-MAP at the lysimeter sites along with the actual field classification parameters from Aqualinc, Van Housen (2015) and Lilburne et al. (2012).

Table A-1 Soil properties

Site	Soil name	Soil class	Re-classed soil	NZTM X	NZTM Y	Average field capacity (Fc)	Average wilting point (WP)	PAW from (FC-WP) (mm)	Saturated moisture content	Saturated hydraulic conductivity (mm/hr)	Weighted average PAW S-MAP	Webb & Carrack soil classification
Christchurch Airport	Rang_5a.1	VL	VL	1561658	5185007	0.159	0.0780	81.0	0.4	>72	64	
Dorie	Temp_1a.1	H	M	1520876	5154624	0.314	0.163	151	0.5	< 4	165.5	Templeton (medium soil)
Dunsandel	Eyre_23a.1	VL	L	1531147	5159702	0.213	0.105	108.0	0.4	4 to 72	73.7	
Hororata	Ruap_2a.1	L	XL	1517008	5180965	0.204	0.130	74.7	0.4	4 to 72	93.9	
Larundel	Paha_2a.1	M	L	1553882	5199527	0.230	0.118	112	0.4	<4	96.1	
Lincoln University	Temp_4a.1	M	H	1556152	5167453	0.338	0.132	205.8	0.5	<4	148.1	

Methven	Mayf_2a.1	L	L	1492110	5162375	0.270	0.140	130.0	0.4	4 to 72	95	Lismore stony silt loam (light soil)
Wakanui	Temp_1a.1	H	D	1502943	5130633	N/A	N/A	N/A	0.4	<4	159.6	Pahau soils (medium and deep soils) PAW 125
Winchmore	Lism_1a.1	L	L	1503008	5150549	0.226	0.122	104.7	0.4	4 to 72	93.8	

Climate and irrigation data

Each Lysimeter site has an associated precipitation record. Lysimeter precipitation rates were compared with the nearest climate station record on NIWA's Cliflo database. NIWA climate station records were used where there was a commonality between rainfall records, or the climate stations were very close to the lysimeters. Preference was given to the climate stations, as these also recorded potential evapotranspiration, and it was assumed they had a more stringent quality assurance program. Lysimeter records were used where the NIWA climate stations were distant, had significant discrepancies in recorded rainfall, or the lysimeter site was irrigated.

Potential evapotranspiration (PET) data was sourced from the nearest climate station recording PET listed in NIWAs CliFlo database (Ashburton Aero station # 319803, 2006 - present, Winchmore station # 318803, 1971 - present, Lismore station # 39845 2012 - present).

For irrigated lysimeters, the irrigation depth increment supplied were either captured in the sites rainfall record or was independently recorded. In all instances where the lysimeter sites were under irrigation, the rainfall and irrigation record was used in parameterising the model.

Methodological approach

Once the source data was compiled, comparisons were made between the modelled and actual recharge data for each lysimeter. The assessment of recharge required matching of both recharge volume and timing of recharge before they were considered reasonable. Where modelled results failed to reproduce reasonable matches to lysimeter recharge, adjustments were made to hydraulic conductivity or field capacity based on S-MAP ranges for the appropriate soils.

Model results

Review of the lysimeter site data suggests that recharge may vary by as much as 20% within a very small area. The high degree of variance reduces confidence that estimates of recharge are representative. Modelling of the various lysimeter sites revealed that the site-specific data was able to produce reasonable approximations of the recharge. No sites needed significant adjustment of parameters to reproduce the observations effectively. In part, this is because of the work Aqualinc had previously done to accurately characterise the soils in a way that was both modellable and consistent with observations. Similarly, this was the case with the soils characterised by Van Housen (2015).

The simple soil moisture balance model can accurately approximate recharge at sites with known soil properties (Figures A-1 through A-12). The ability to reproduce observations at the various lysimeter sites with little alteration to previous estimates of soil parameters provides confidence in the functioning of the modelling platform and the adopted approach. It is, therefore concluded that the model could be used to produce a reasonable approximation of recharge in other areas (without lysimeter data) with similar soils.

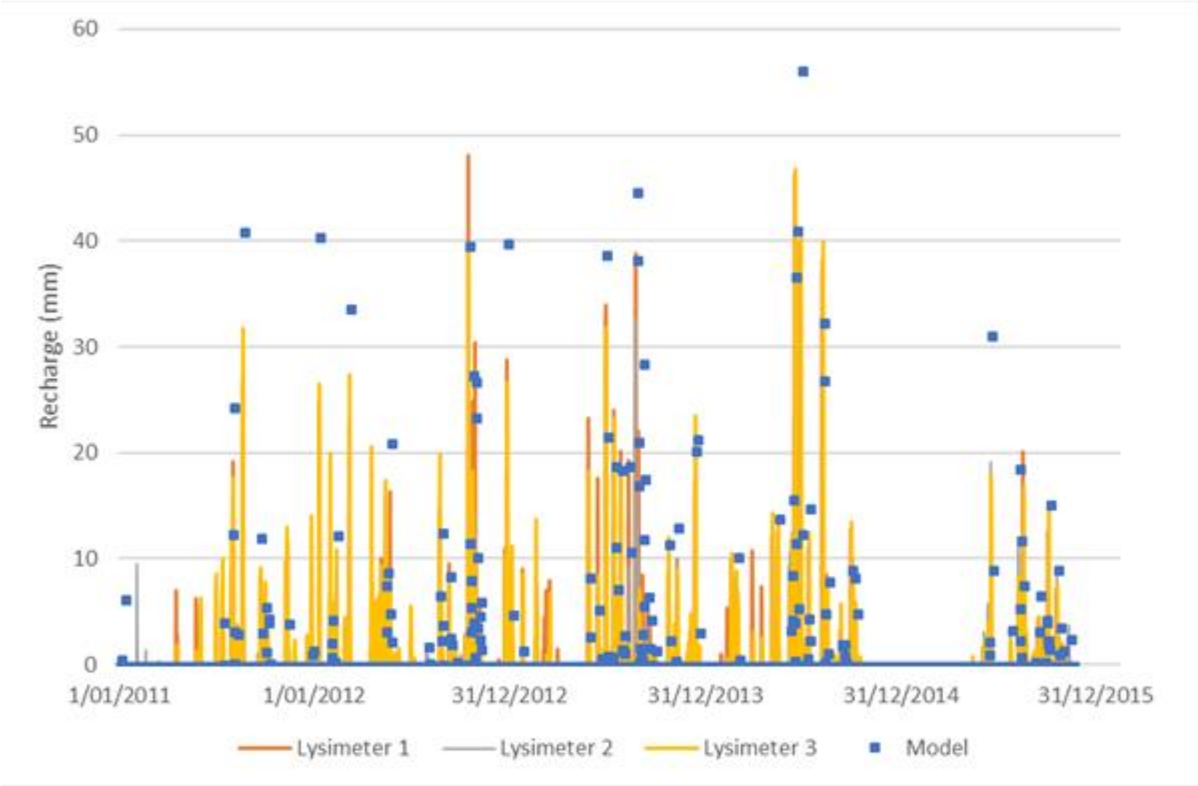

Figure A-1 Methven irrigated lysimeters incremental recharge.

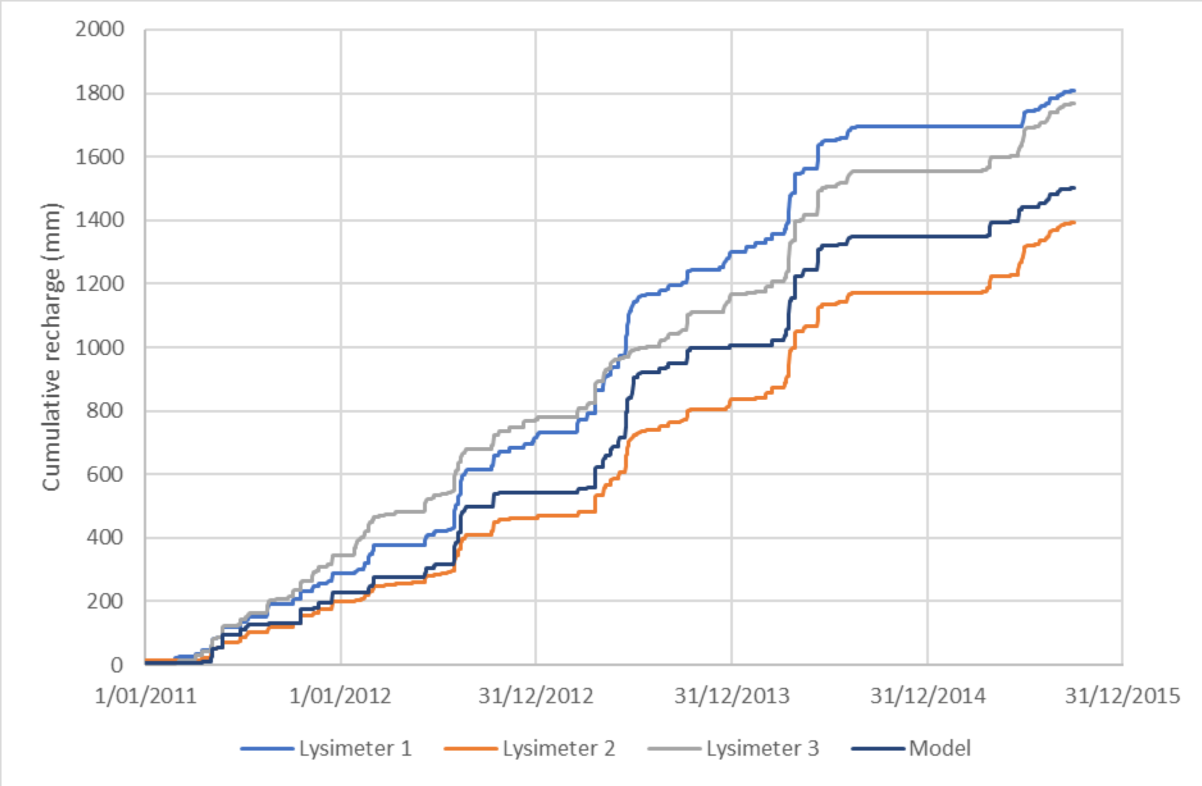

Figure A-2 Methven irrigated lysimeters cumulative recharge.

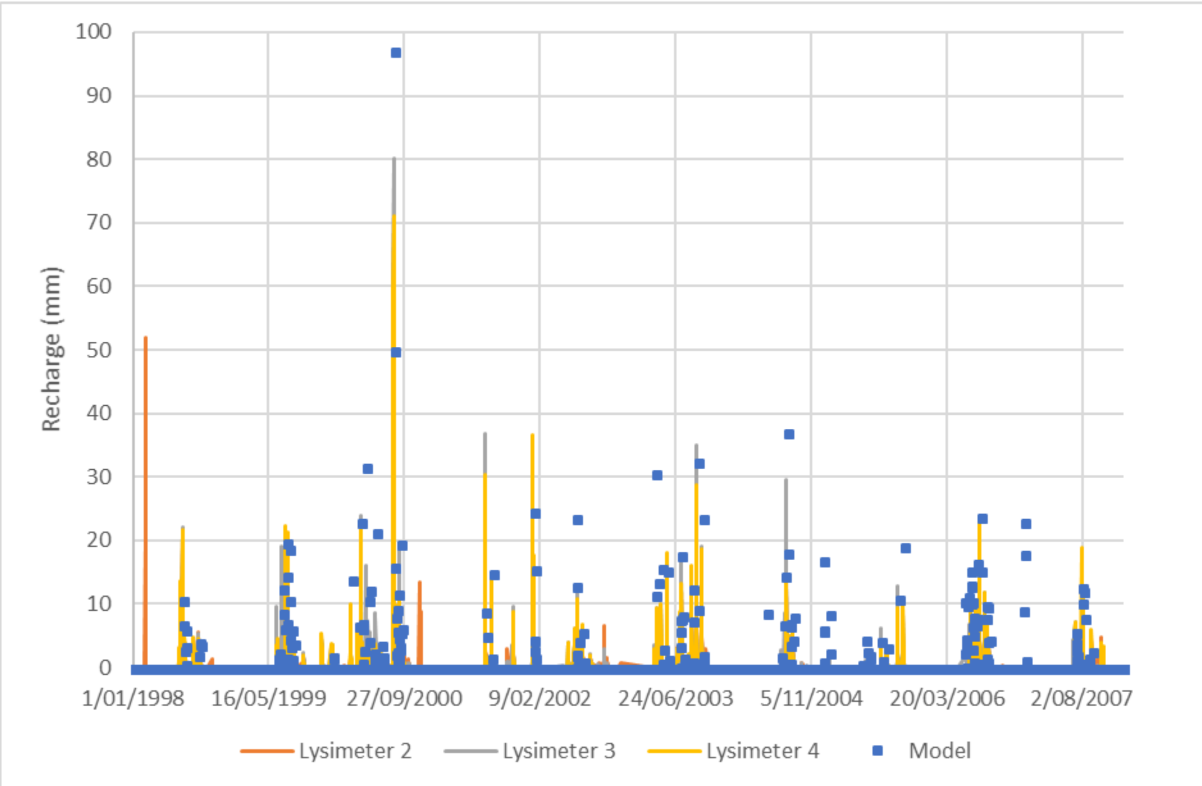

Figure A-3 Winchmore dryland lysimeters incremental recharge.

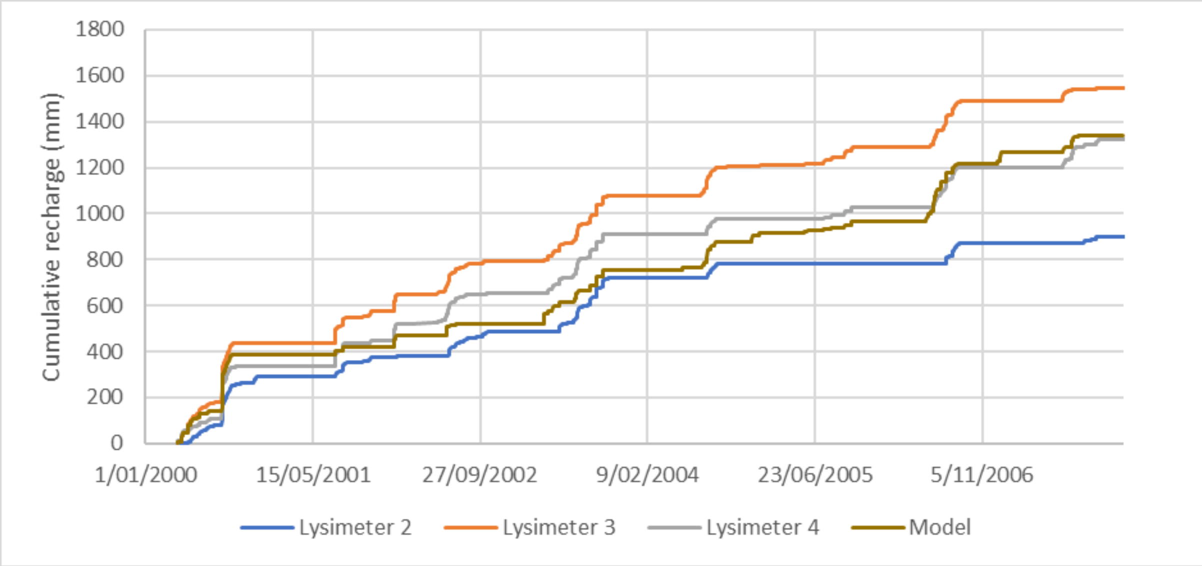

Figure A-4 Winchmore dryland lysimeters cumulative recharge.

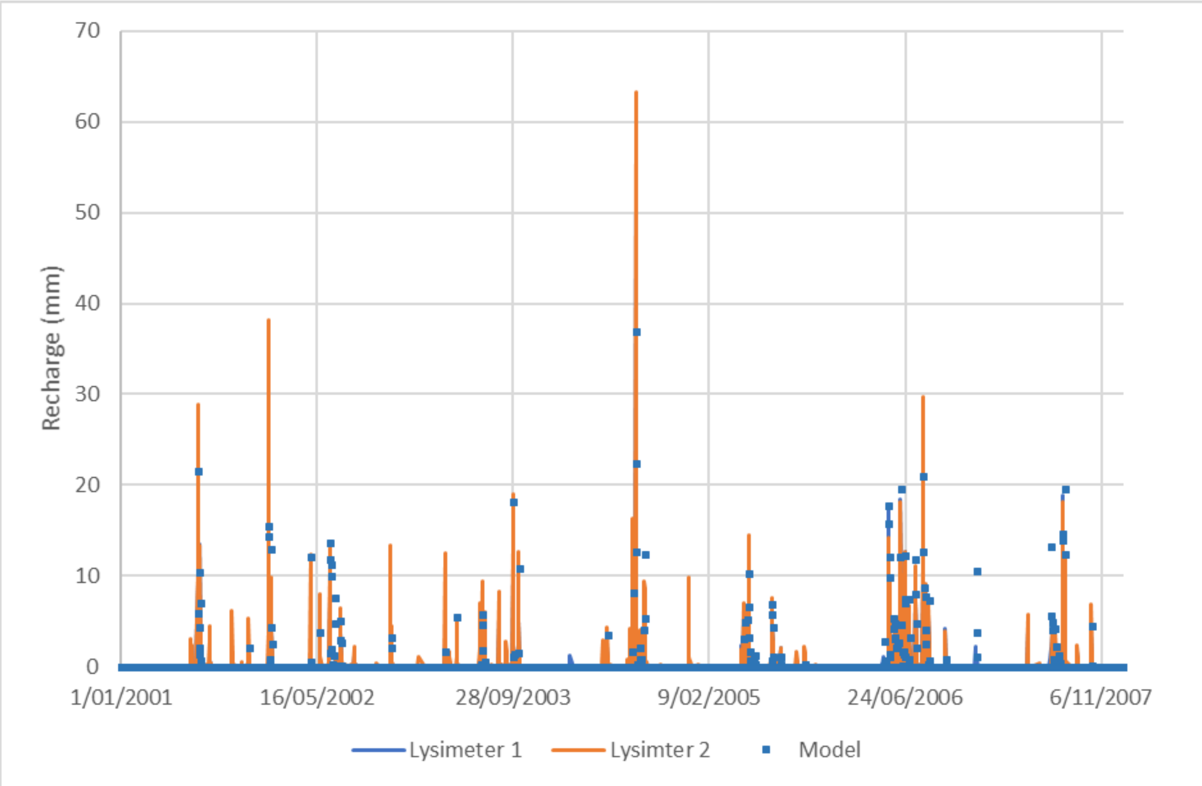

Figure A-5 Christchurch Airport dryland lysimeters incremental recharge.

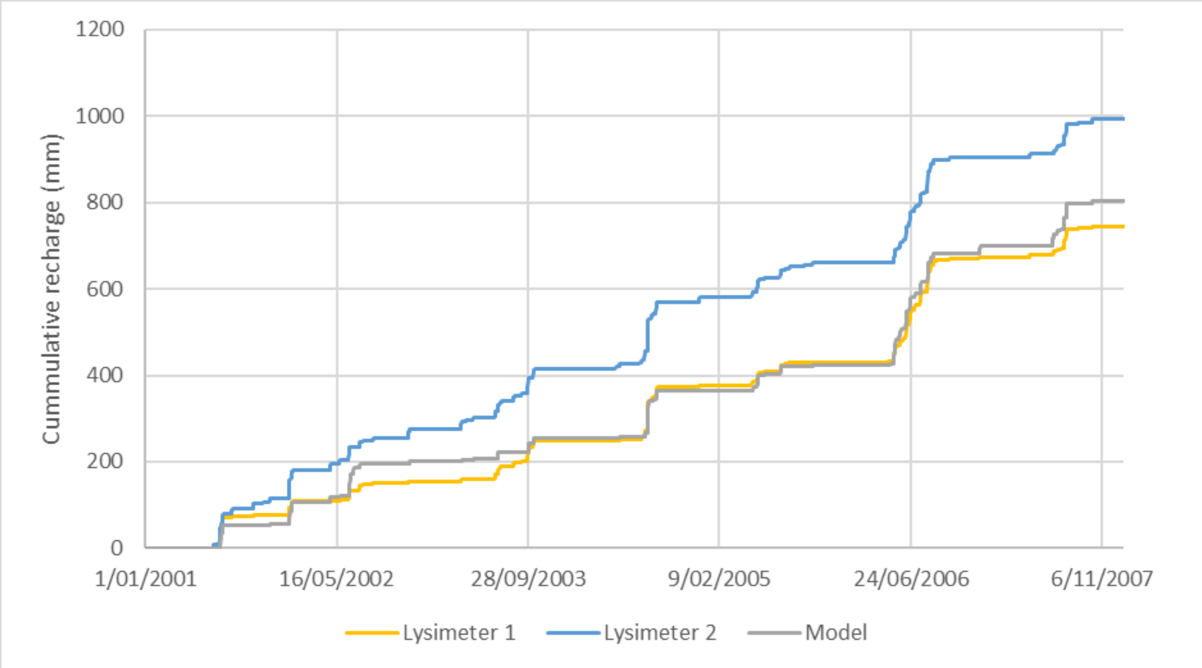

Figure A-6 Christchurch Airport dryland lysimeters cumulative recharge.

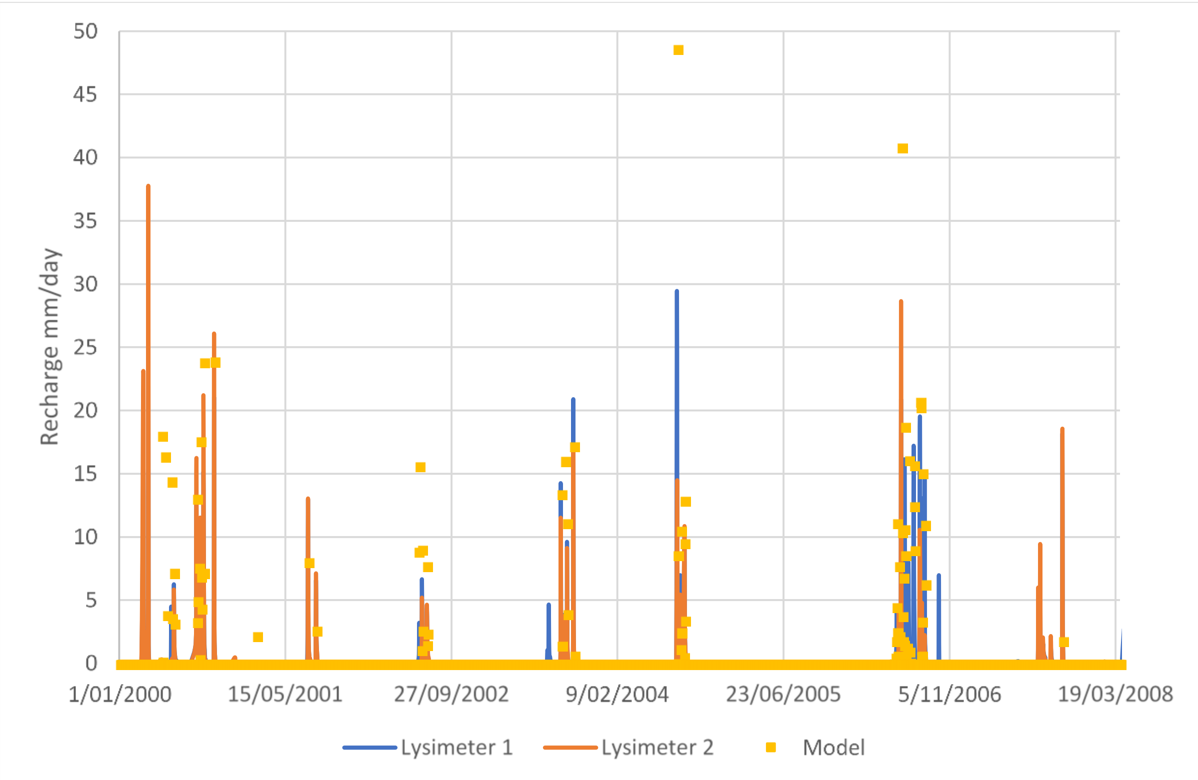

Figure A-7 Lincoln dryland lysimeters incremental recharge.

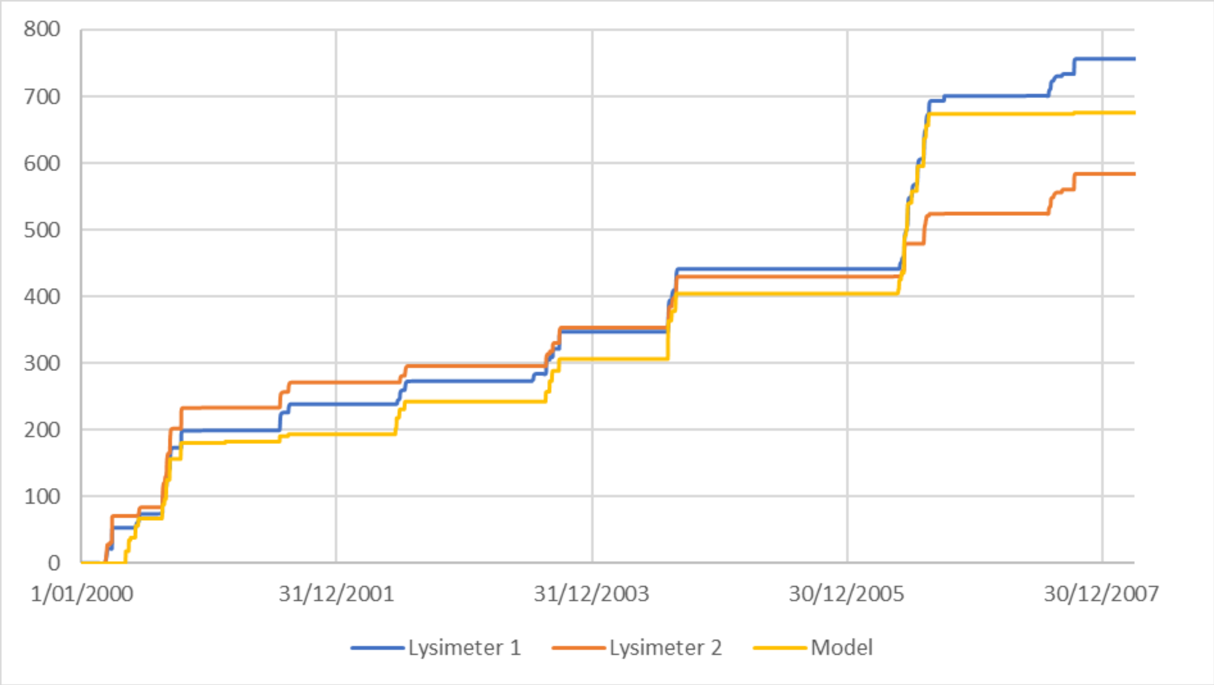

Figure A-8 Lincoln dryland lysimeters cumulative recharge.

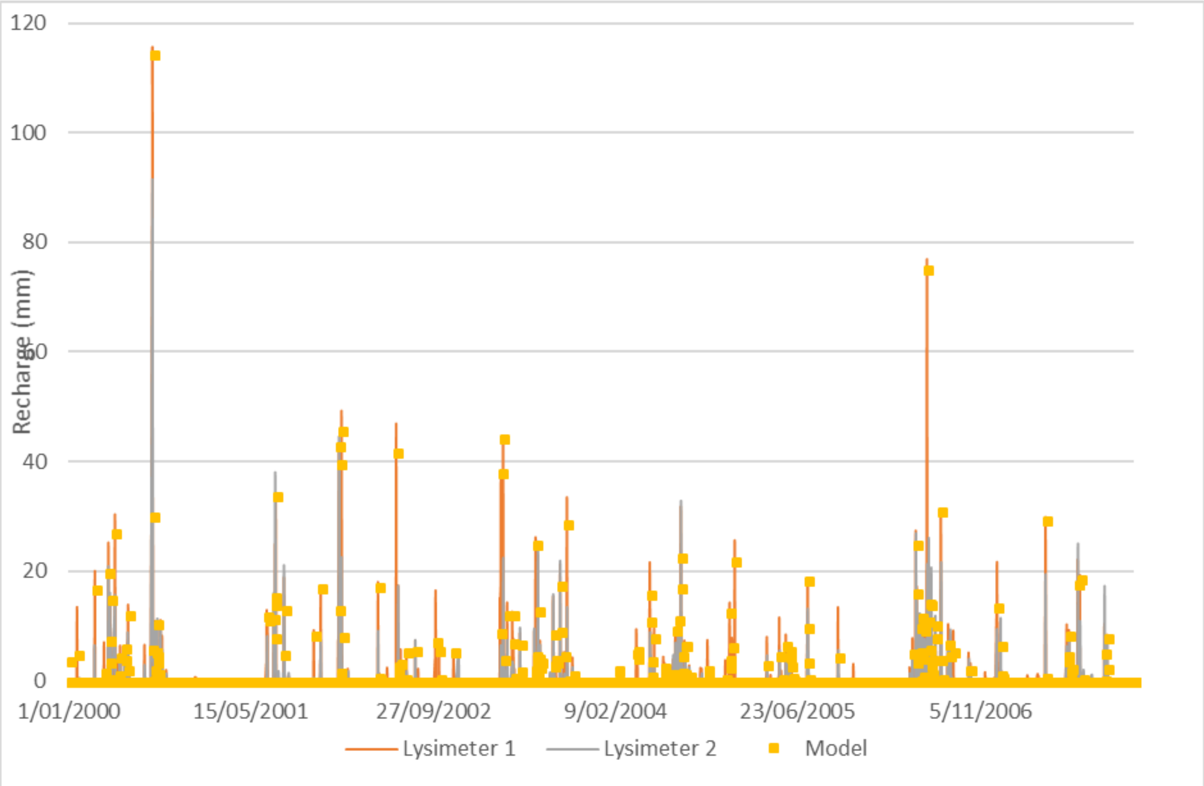

Figure A-9 Hororata dryland lysimeters incremental recharge.

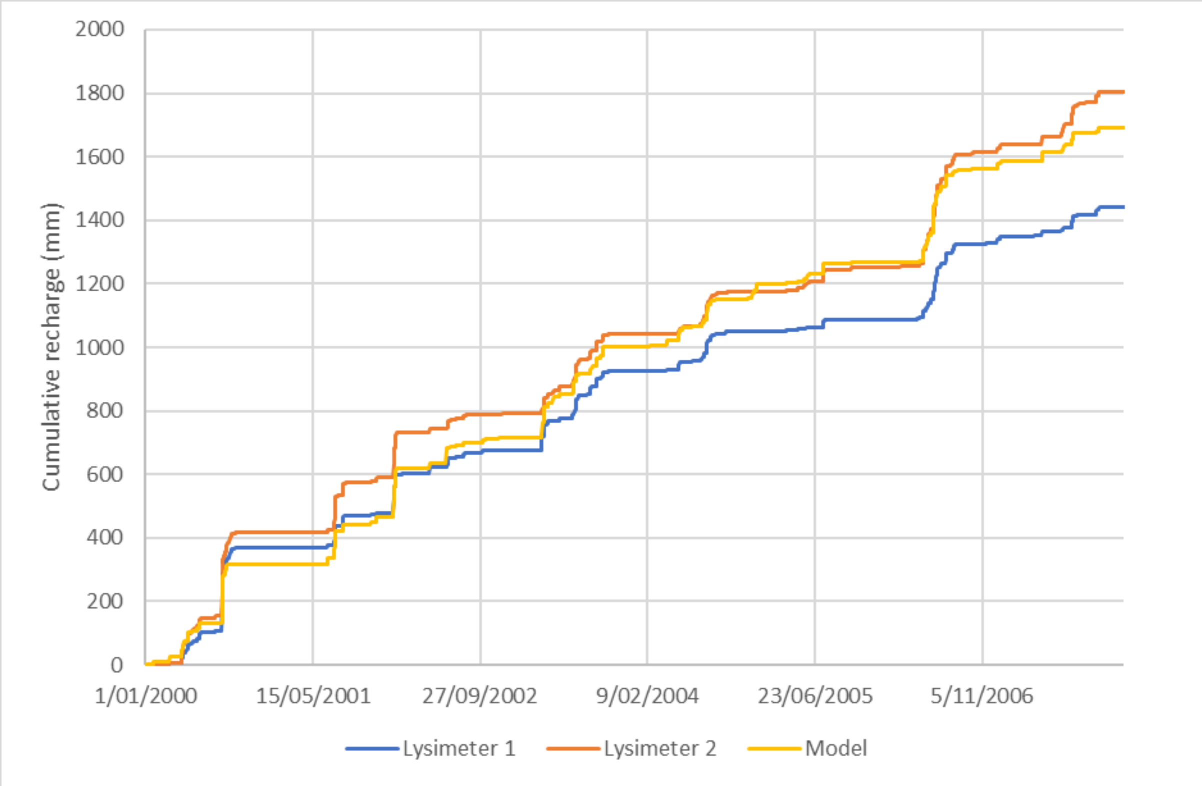

Figure A-10 Hororata dryland lysimeters cumulative recharge.

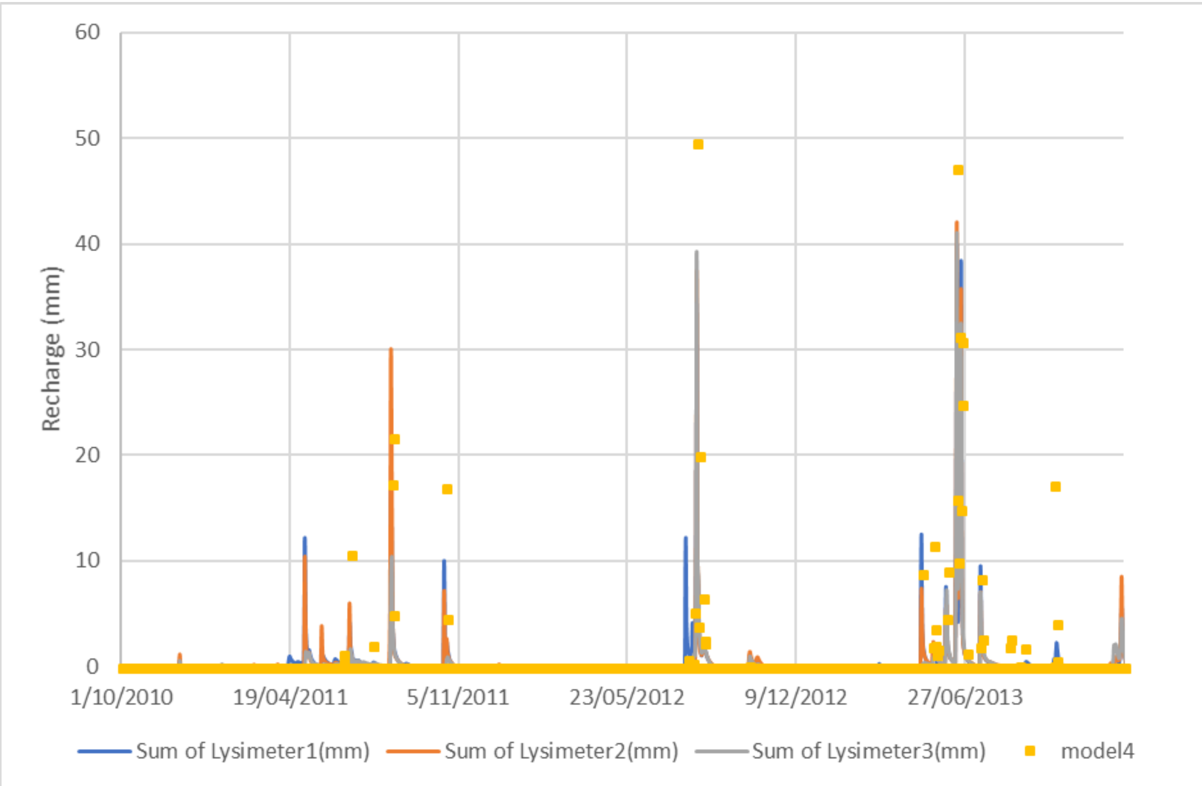

Figure A-11 Dorie irrigated lysimeters incremental recharge.

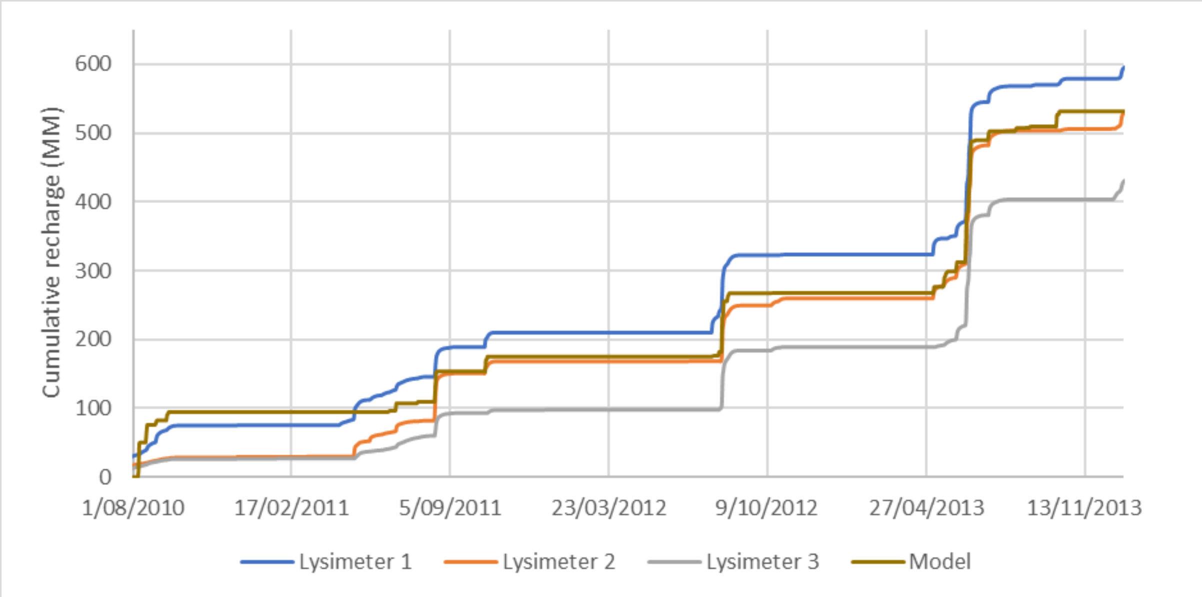

Figure A-12 Dorie irrigated lysimeters cumulative recharge.

Limitations of the soil moisture balance model

Problems with the soil moisture balance approach, to estimating recharge, are encountered when the discrepancies between S-MAP and local site characterisation are considered. The data reveals considerable discrepancies between the 250,000 scale S-MAP viewer (and Arc GIS files) and the site-specific soil descriptions.

Additionally, when the S-MAP soil fact cards for a specific soil are reviewed, there is a further discrepancy with the S-MAP viewer. The discrepancies present a problem for the estimation of recharge at sites without dedicated field mapping by qualified soil scientists. The discrepancies show that not only does S-MAP fail to map the general soils in many instances accurately, but it also estimates different profile available water (PAW) values than the associated soil fact cards.

Alternative recharge approaches

The possibility of using a chloride mass balance approach for estimating recharge was investigated, however, due to lack of data and heavy groundwater irrigation across the Hinds Plains, the approach was deemed unfeasible and was not pursued.

Natural groundwater level fluctuations present another approach to the estimation of aquifer recharge. Using this approach, an increase in groundwater level elevation is proportional to recharge and specific yield of the aquifer. Initially, it was anticipated that the water level fluctuations in well K37/1748 could be used to estimate rainfall recharge near the MAR trial site. However, a review of the water levels in well K37/1748 revealed the well responded more to the presence or lack thereof of water in a stock water distribution race immediately adjacent to it. As loss rates from this reach of the stock water distribution race were unknown, as was the operational timing of the water distribution, the approach was ruled out as providing a useful estimate of recharge.

Model domain recharge layer

The soil properties calibrated to the lysimeter sites have been used to produce an average recharge layer for the model domain. S-MAP (Landcare Research 2014) was used to provide soil coverage details (Figure A-13), while an irrigated land-use layer from ECAN (Figure A-14) was used to determine the areas that were irrigated. Theisen polygons were used to describe precipitation and PET inputs (Figure A-15 and A-16). The final resulting average groundwater recharge layer is shown in Figure A-17.

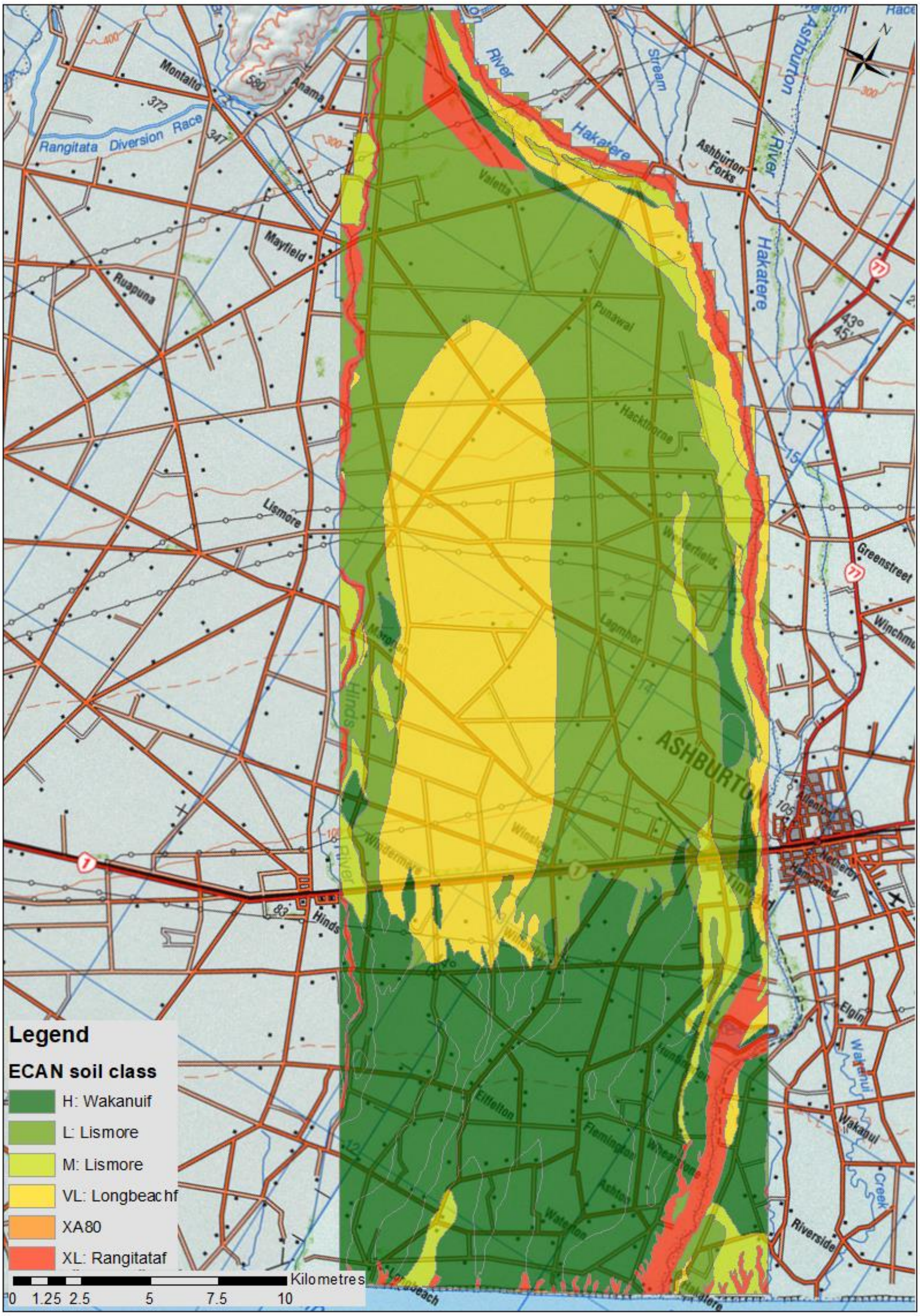

Figure A-13 Soil Class (S-MAP) (Landcare research).

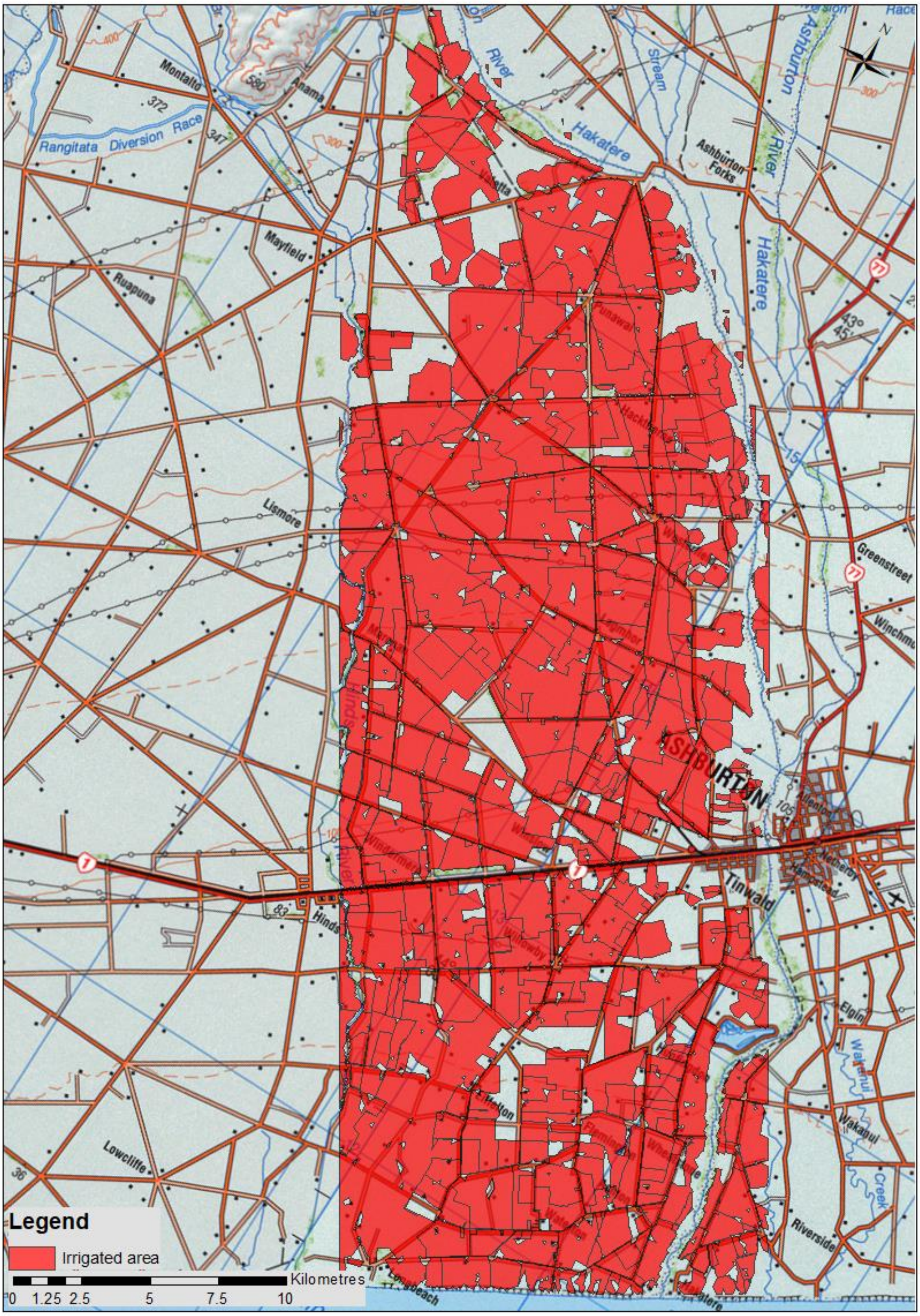

Figure A-14 Irrigated area.

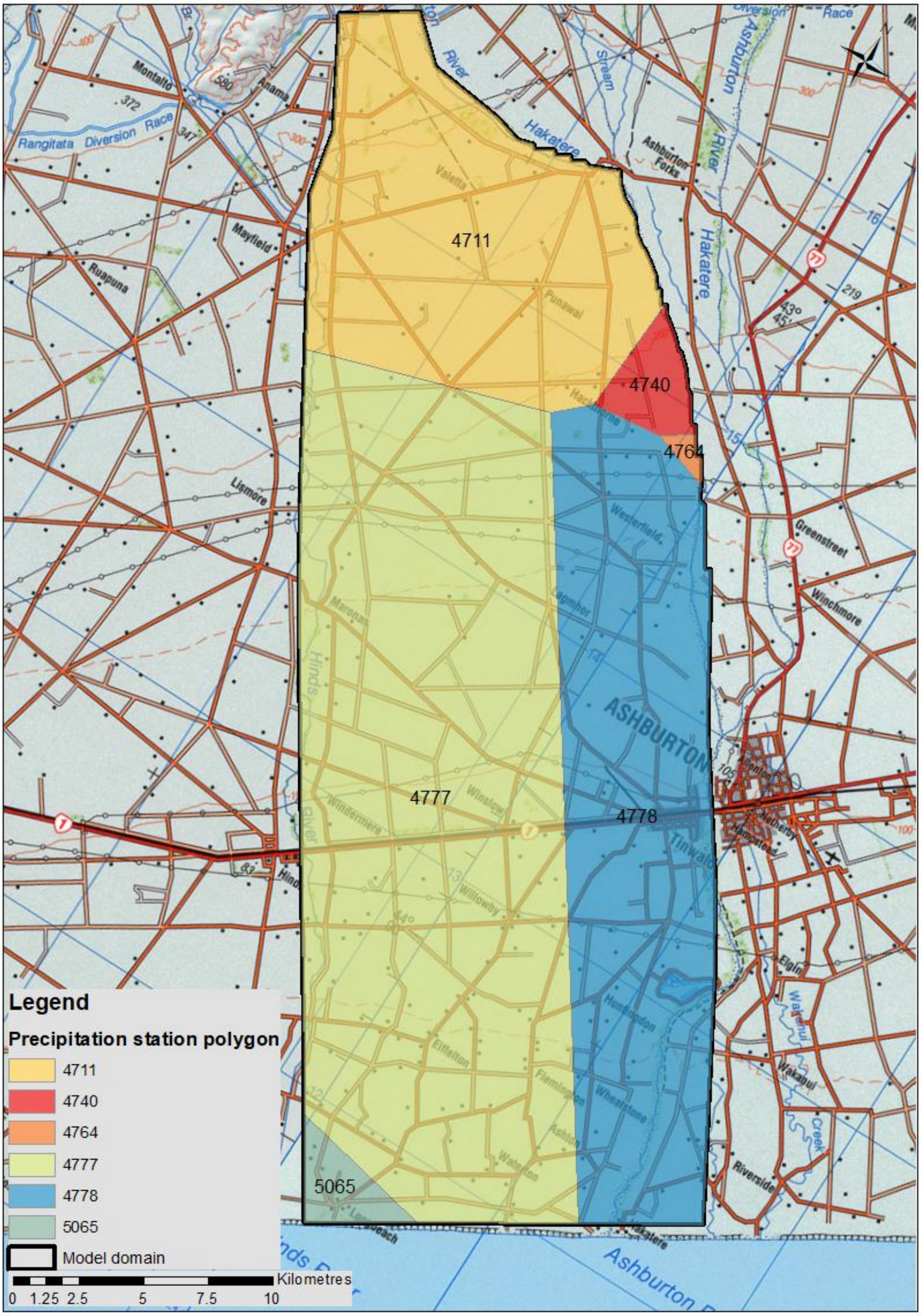

Figure A-15 Thiessen polygons of precipitation climate stations.

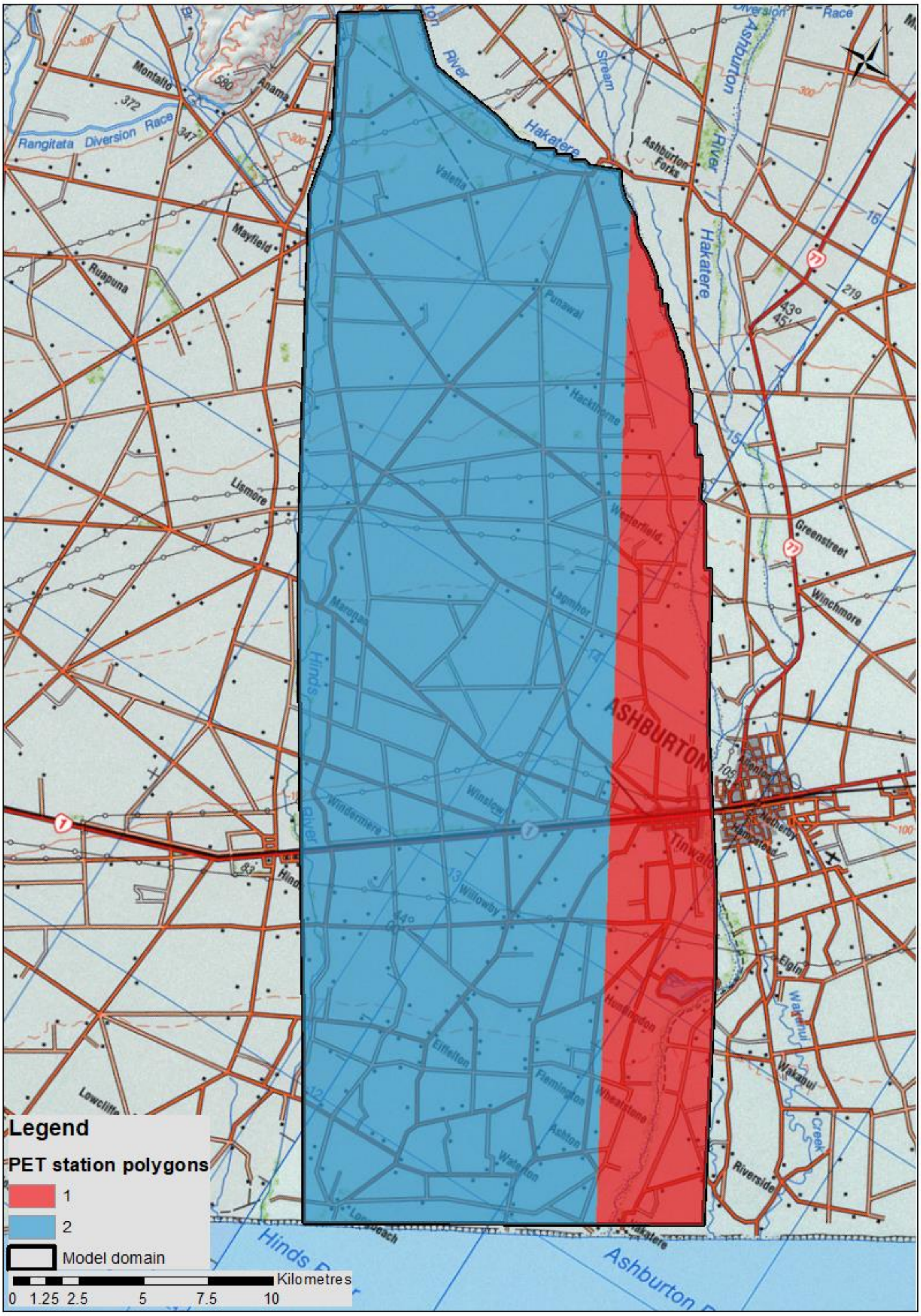

Figure A-16 Potential evapotranspiration station thiessen polygons.

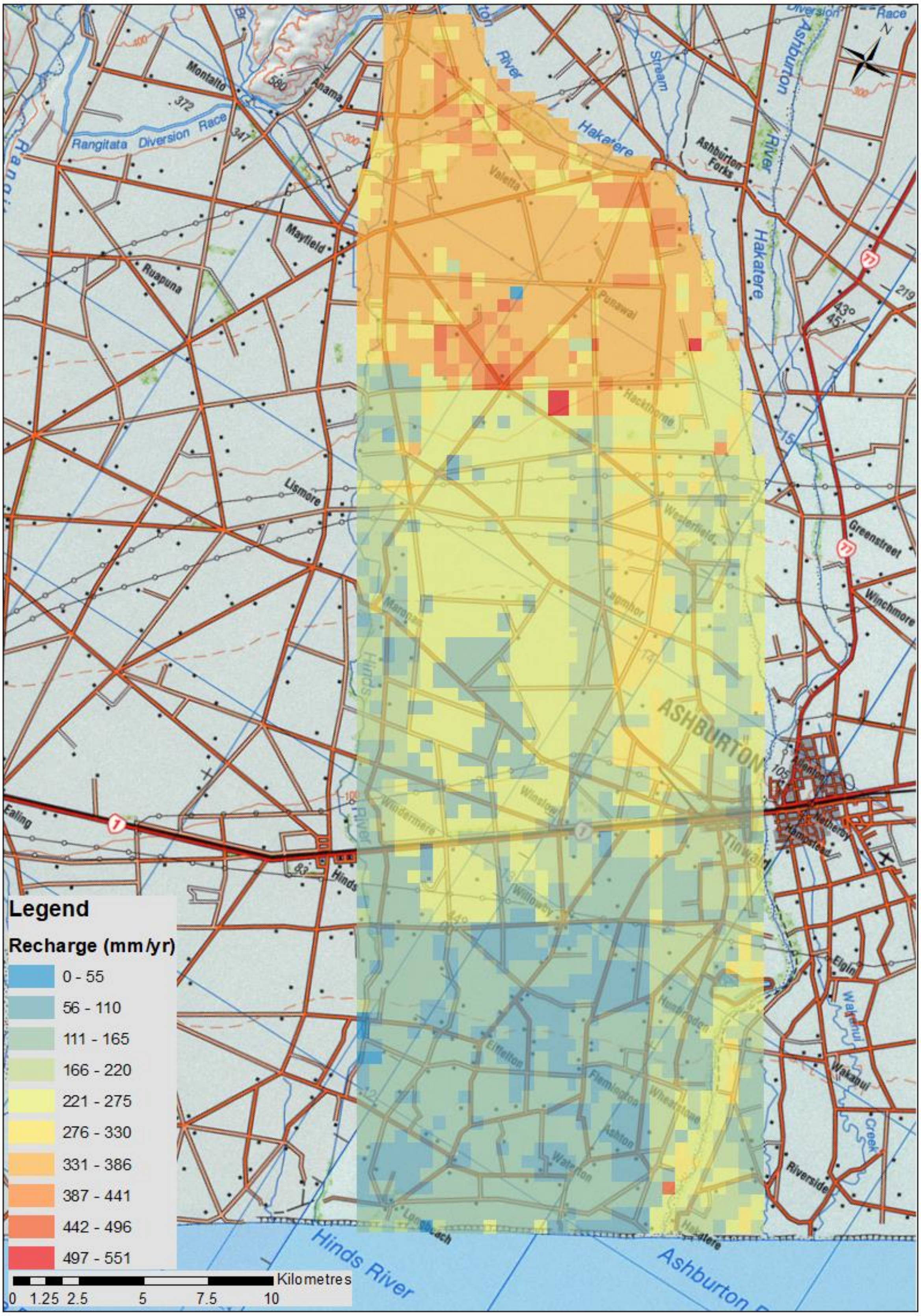

Figure A-17 Average groundwater recharge.

The simple soil moisture balance model has been produced for the primary soils encountered in the Ashburton region. The model can replicate measured recharge at lysimeter sites across Canterbury. Problems with the regional scale inputs (e.g. S-MAP) cast doubt over the sole use of soil-based recharge models for estimating groundwater recharge. However, presently, the author considers the soil moisture balance approach the most robust, considering the confounding factors impacting the other potential methods.

Appendix B Analytical analysis of mounding

The water level fluctuations caused by barometric efficiency need to be removed for the actual water level fluctuations to be accurately quantified. Therefore, the first step in data analysis is to assess barometric efficiency. The observation wells used in this assessment were measured with non-vented pressure transducers that measure absolute pressure, including atmospheric. Assessment of barometric efficiency requires groundwater level observations of pumping or recharge. To this end, measurements are needed before or following the periods of pumping or recharge. Observations were collected before the commencement of the trial and analysed to assess barometric efficiency. An example of the barometric efficiency correction is presented in Figure B-1 and Figure B-2. The barometric efficiency in the observation wells (close to 100%) generally supports the unconfined aquifer conceptual model.

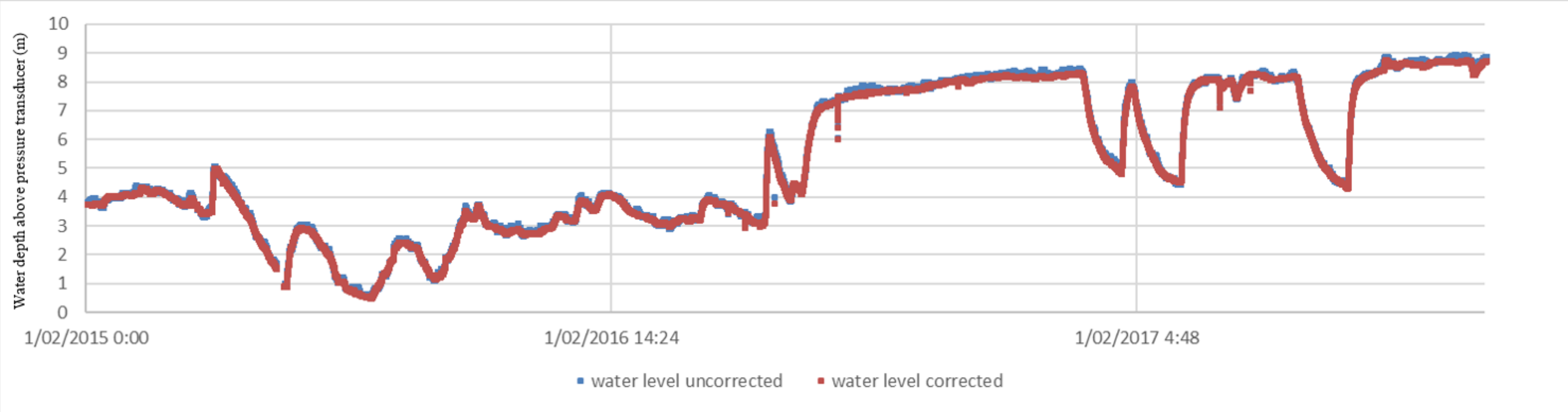

Figure B-1 Barometric efficiency correction well K37/1748.

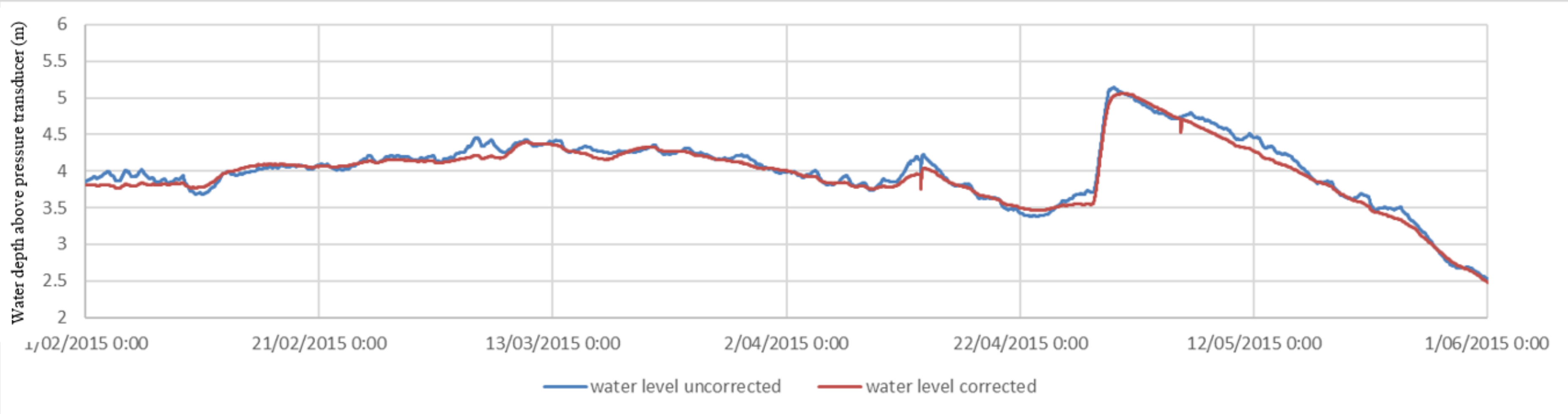

Figure B-2 Barometric efficiency correction for well K37/1748, zoomed in on a short period.

Analysis of mounding response

Aqtesolv™ has been used to analyse the trial response. Aqtesolv™ is an aquifer test analysis software developed by HydroSOLVE Inc. (2007). Initial modelling has focused on assessing the response using the Cooper-Jacob (1946) method. Cooper-Jacob is a straight-line approximation of the Theis solution and assumes a fully penetrating well in confined aquifer conditions. The solution may be applied to unconfined conditions by correcting drawdown, using the formula $s' = s - s^2/2b$, where s is drawdown, b is the aquifer thickness and s' is corrected drawdown. Cooper-Jacob is useful as a first pass for estimating transmissivity from early time-drawdown data. In the author's experience, due to the nature of Canterbury's gravel aquifers, it becomes less useful for later time-drawdown when other factors, such as leakage of groundwater from overlying strata occur. To analyse the mounding response, the aquifer thickness is assumed to be 30 m, the depth below which, ECAN database of aquifer tests suggests the aquifer generally ceases to behave in a way that resembles unconfined conditions.

Analysis has focused on the wells showing the clearest response to the recharge trial, specifically well K37/1748 and BY20/0149. K37/1748 is located immediately down-gradient from the MAR trial site. Barometric efficiency is close to 100% for this well, supporting the unconfined hypothesis. Water levels are higher than the surrounding wells due to the presence of a stock water supply race that runs directly past the well. Both well K37/1748 and well BY20/0149 are potentially too close to the trial pond, thereby reducing confidence that can be placed in the data.

The early time response to recharge is variable in wells K37/1748 and BY20/0149, leading to Golder (2017) attributing early-time drawdown to the Lesse effect (air entrapment). However, when Barometric corrections are applied to the data, it appears that the early-time drawdown may be in response to barometric pressure and not air entrapment.

Set-up of the analytical model involved replication of the trial set up, including estimates of losses along with the race supplying water to the recharge basin. Multiple pumping wells were used to simulate the recharge from the distribution race and recharge basin.

Initial estimates of aquifer parameters using the Cooper-Jacob straight-line solution are presented in Figure B-3 and Figure B-4. As can be seen, depending on the slope analysed, the resulting parameters are by no means unique, with multiple transmissivities possible to estimate. Figure B-3 and Figure B-4 represent initial estimates of the upper and lower bounds for transmissivity.

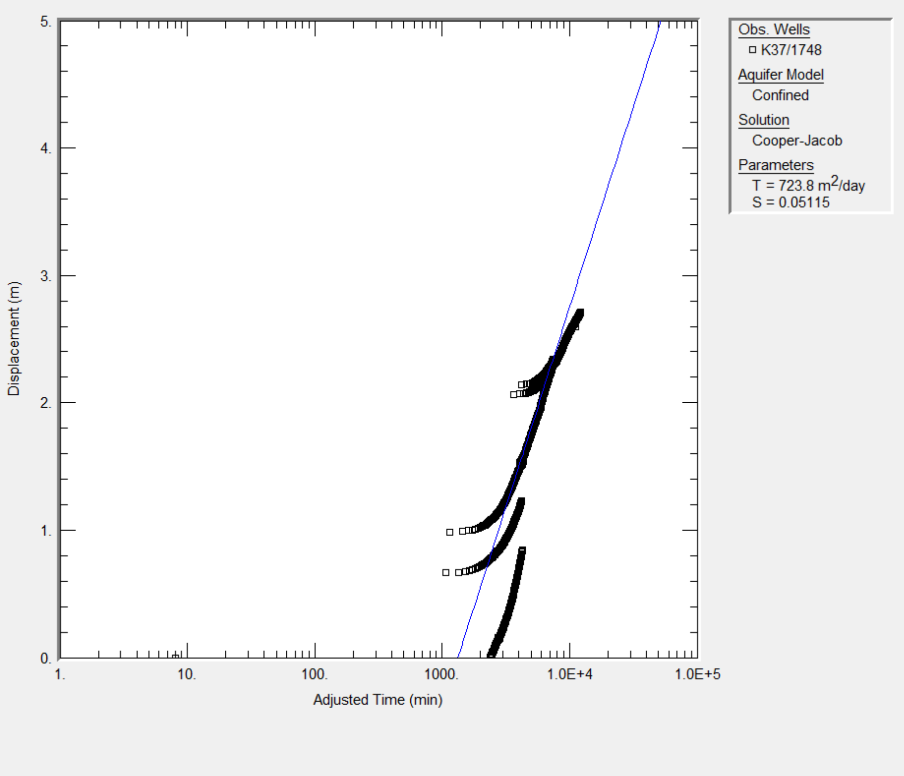

Figure B-3 Well K37/1748 Cooper-Jacob early time drawdown estimates of upper bounds of aquifer transmissivity.

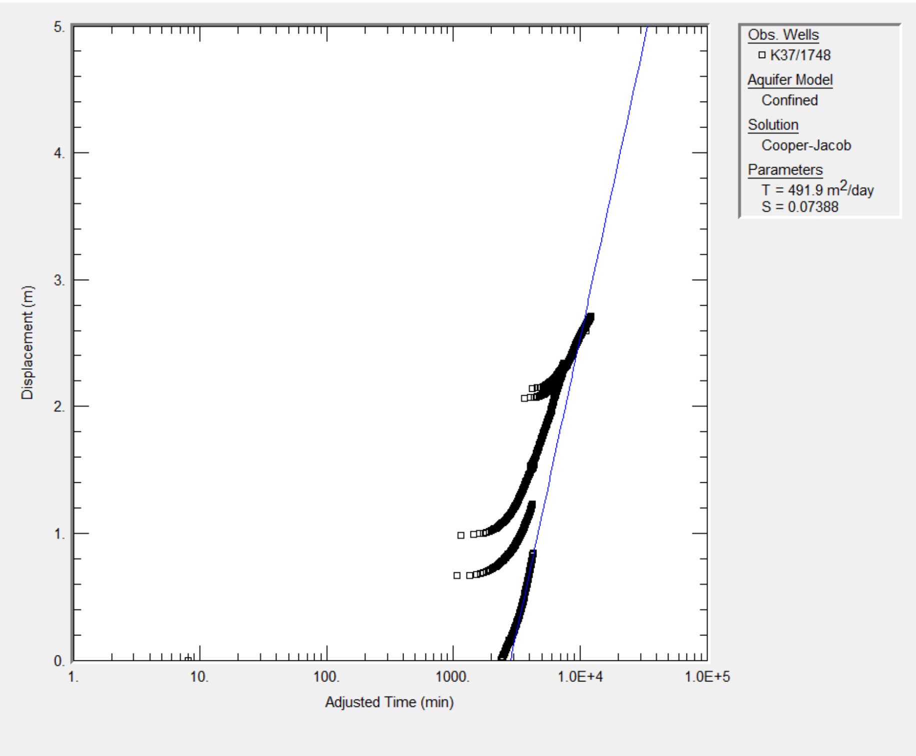

Figure B-4 Well K37/1748 Cooper-Jacob early time drawdown estimate of lower bounds of aquifer transmissivity.

Following the first pass estimate of aquifer parameters, the use of more advanced analytical solutions was investigated to assess the possible aquifer parameters near the site. These more advanced analytical solutions, ranging from the Theis unconfined to the Neuman-Witherspoon (1969) leaky solution, reveal that the problem is more complex than the assumptions that underlie traditional analytical aquifer test solutions.

Accepting the conceptual model of an unconfined aquifer, supported by the barometric efficiency of the observation wells, the Theis method of history matching observations was investigated. Despite the complexity of the model set-up, for the Theis equation to come close to matching the observed data using the transmissivity estimates from Cooper-Jacob straight-line method, the estimated storativity required is beyond what is reasonable for the geology and approaching that of 0.4 (Figure B-5). Various alternative transmissivity ranges were also used, but in general, specific yield estimates were too high. This result implies that either the conceptual model is faulted or that the problem is too complex for the analytical solutions to produce meaningful results and parameter estimates.

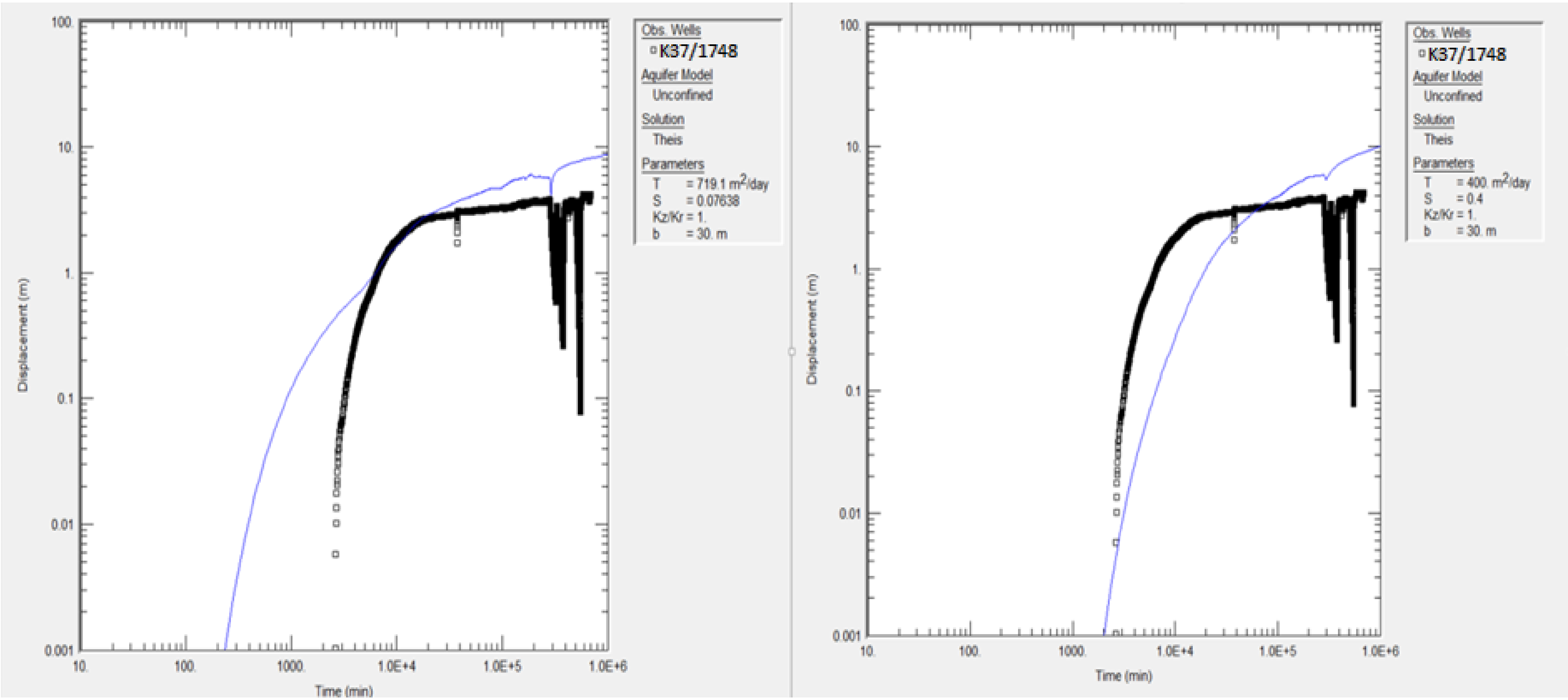

Figure B-5 Theis solution for well K37/1748.

As an attempt to proceed with the analytical approach to estimating the aquifer's parameters, the conceptualisation of the aquifer was changed to assume semiconfined conditions. The Neuman-Witherspoon (1969) semi-confined solution was employed for analysis (Neuman and Witherspoon 1969). Again, despite the introduction of several further parameters such as leakage, overlying transmissivity and storage, the same problem was encountered; namely that the specific yield estimated from the observation wells was too high (0.4-0.6 when matched to BY20/0149) when both wells of the same depth were considered to be in the same aquifer. By remodelling the data excluding the first day of the trial where no response was observed, it is possible to produce a more reasonable estimate (matching K37/1748) at 0.02 (Figure B-6). The recovery response in Figure B-6 is caused by recharge being temporarily shut down.

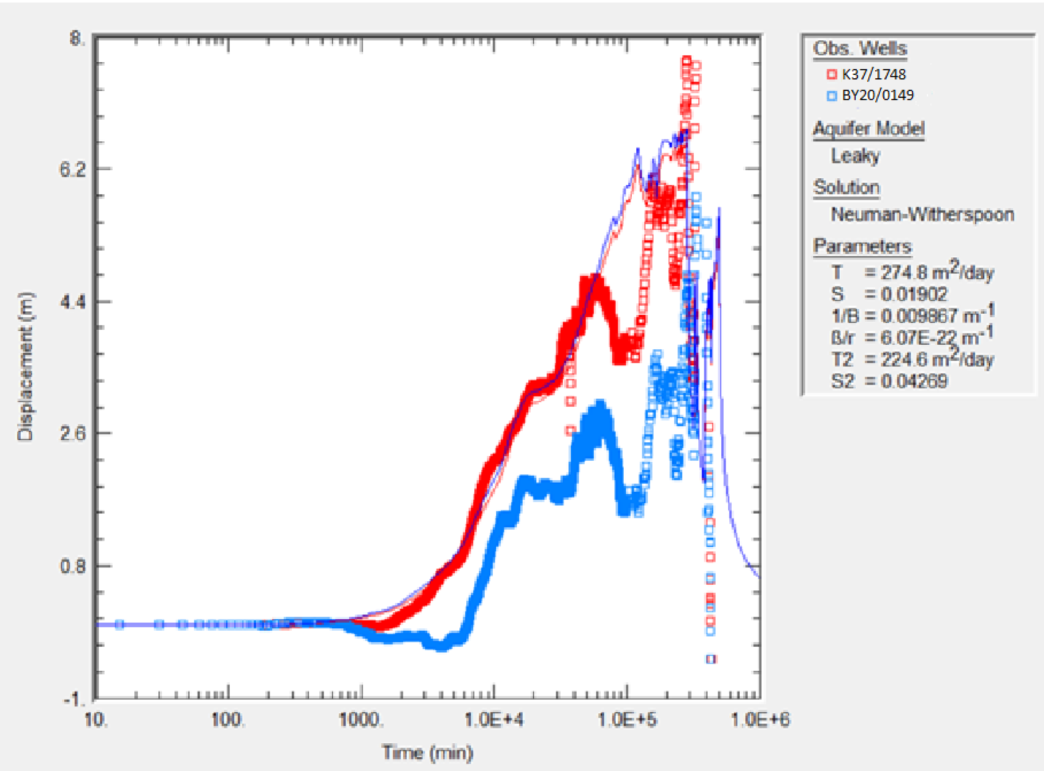

Figure B-6 Neuman-Witherspoon solution, highlighting the inability to match specific yield in both wells installed to the same depth. Modelled with both wells in the second unpumped aquifer.

An investigation into the possible causes for the erroneous storage values led to a review of the well driller's comments; these show that the water table was encountered near the bottom casing in well BY20/0149, implying several tens of meters of unsaturated zone. The presence of unsaturated zone is enough to cause the delay in observed drawdown (mounding), i.e., the delay is caused by the time it took for the recharge water to percolate to the water table.

Knowing that the delayed drawdown (mounding) response is caused by unsaturated flow, it is still possible to use an analytical solution to investigate transmissivity and specific yield/storativity of the aquifer. Using the secondary transmissivity and storativity of the Neuman-Witherspoon (1969) leaky aquifer solution, it is possible to estimate the properties of the aquifer; this does, however, require accepting that the estimates of primary transmissivity, storativity and leakage are compensatory parameters required to account for vertical flow through the initially unsaturated parts of the aquifer (Figure B-7).

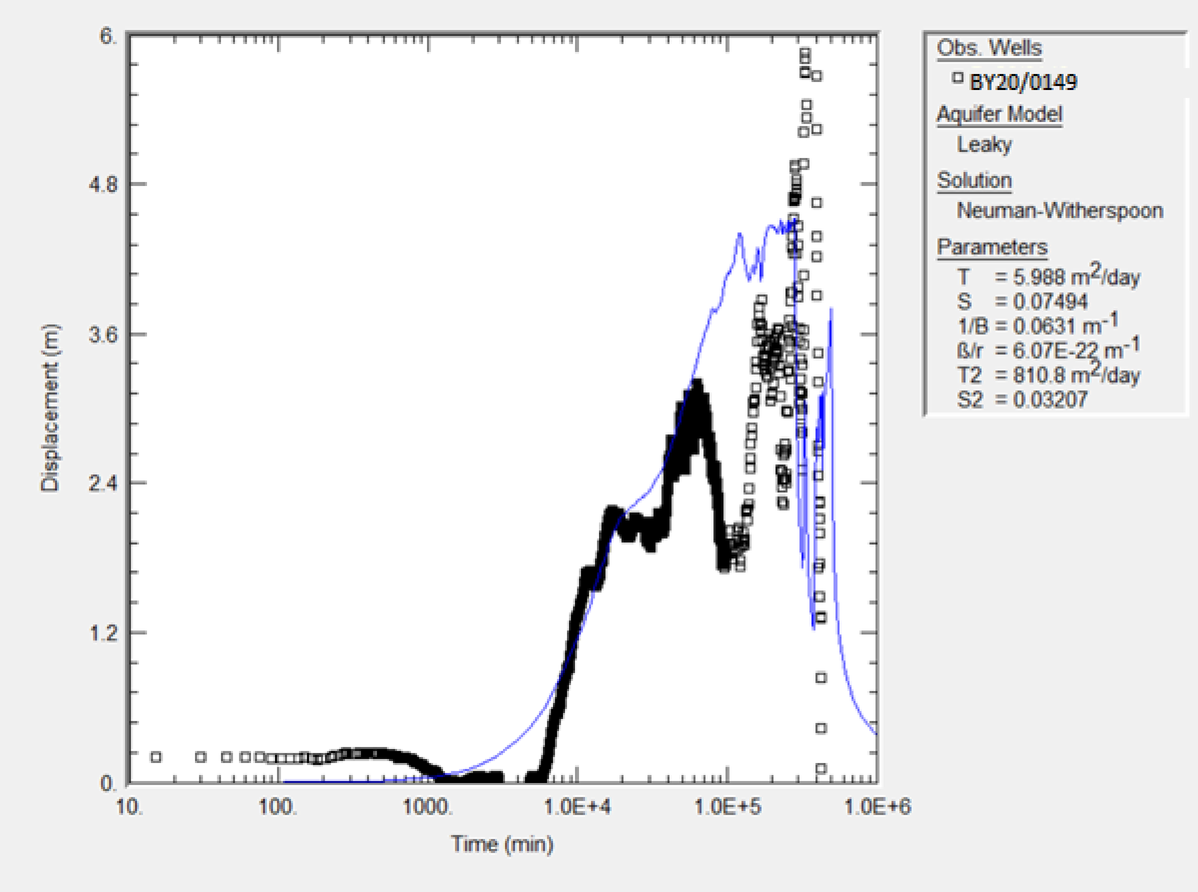

Figure B-7 Neuman-Witherspoon analysis using the overlying aquifer properties as compensatory parameters.

Appendix C Model layer structure

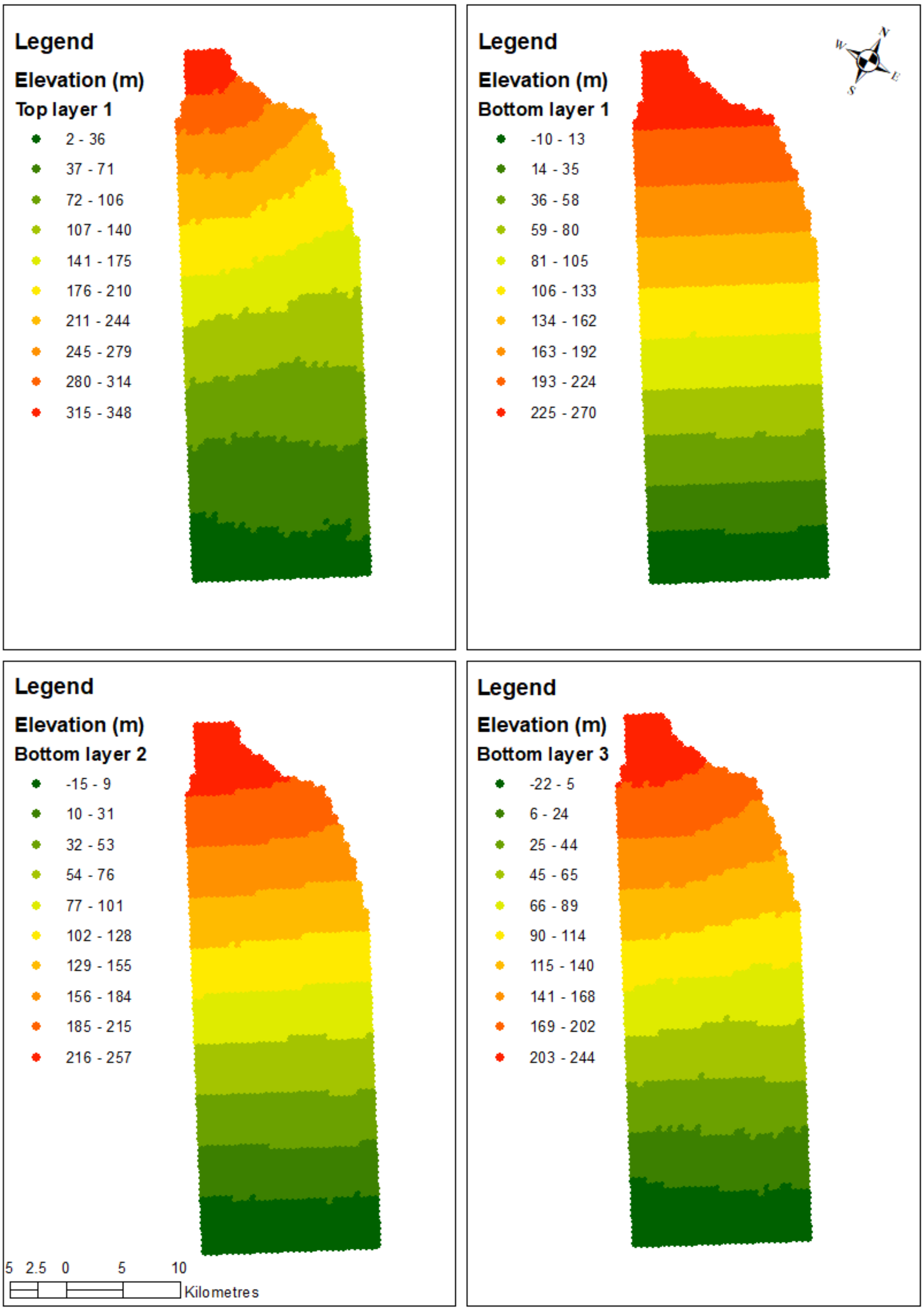

Figure C-1 Layer 1 to Layer 3 elevations.

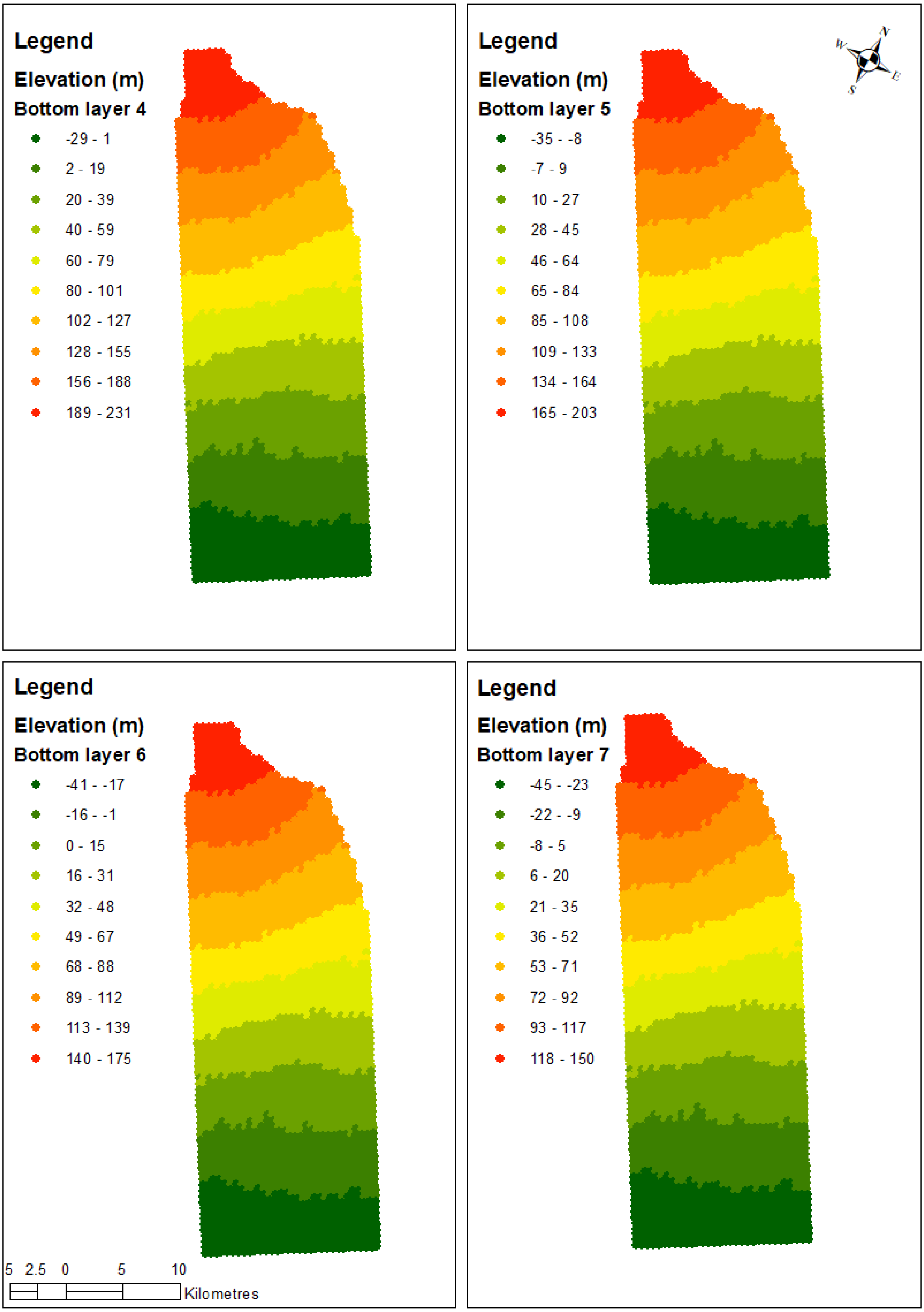

Figure C-2 Layer 4 to layer 7 elevations.

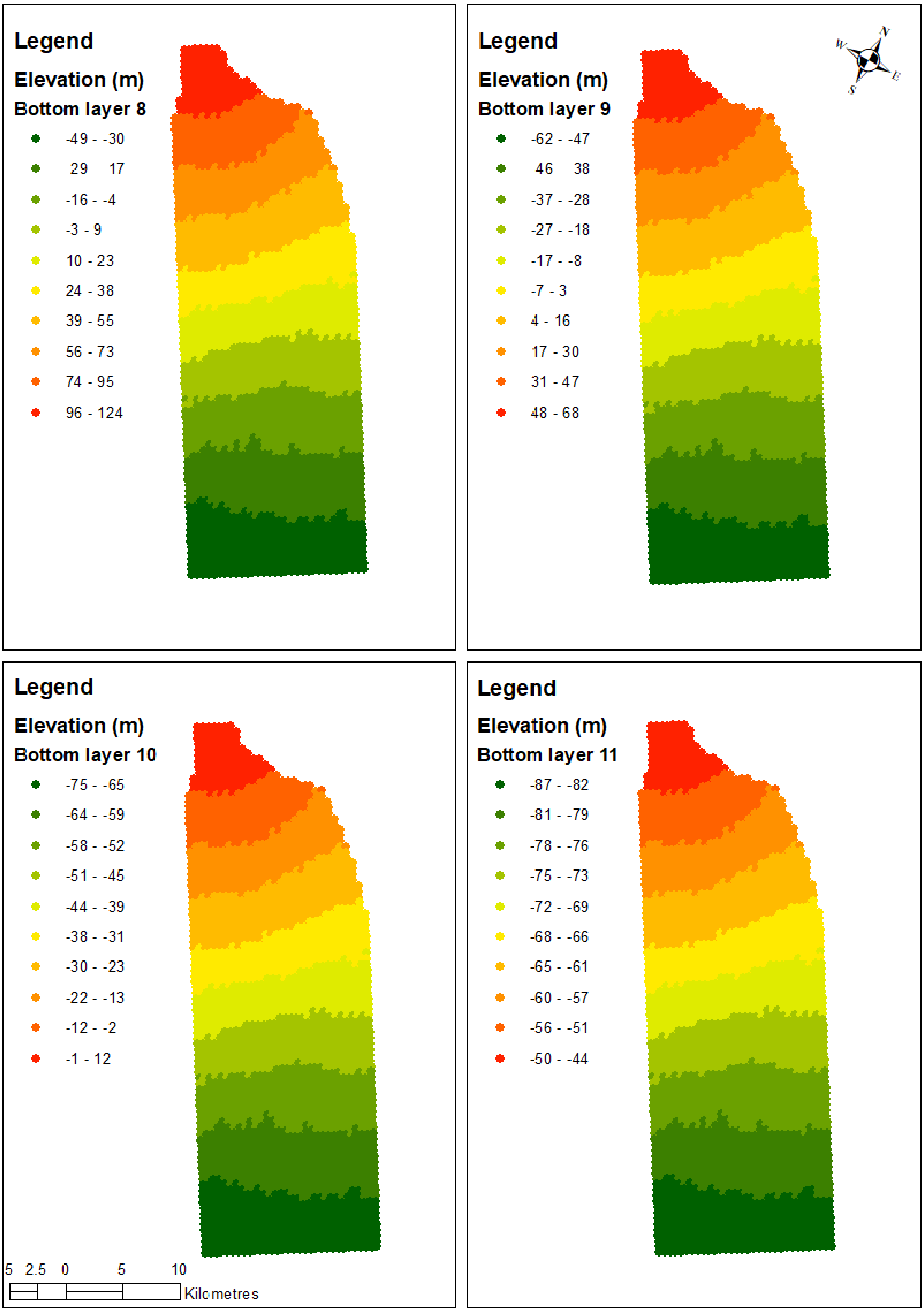

Figure C-3 Layer 8 to layer 11 elevations.

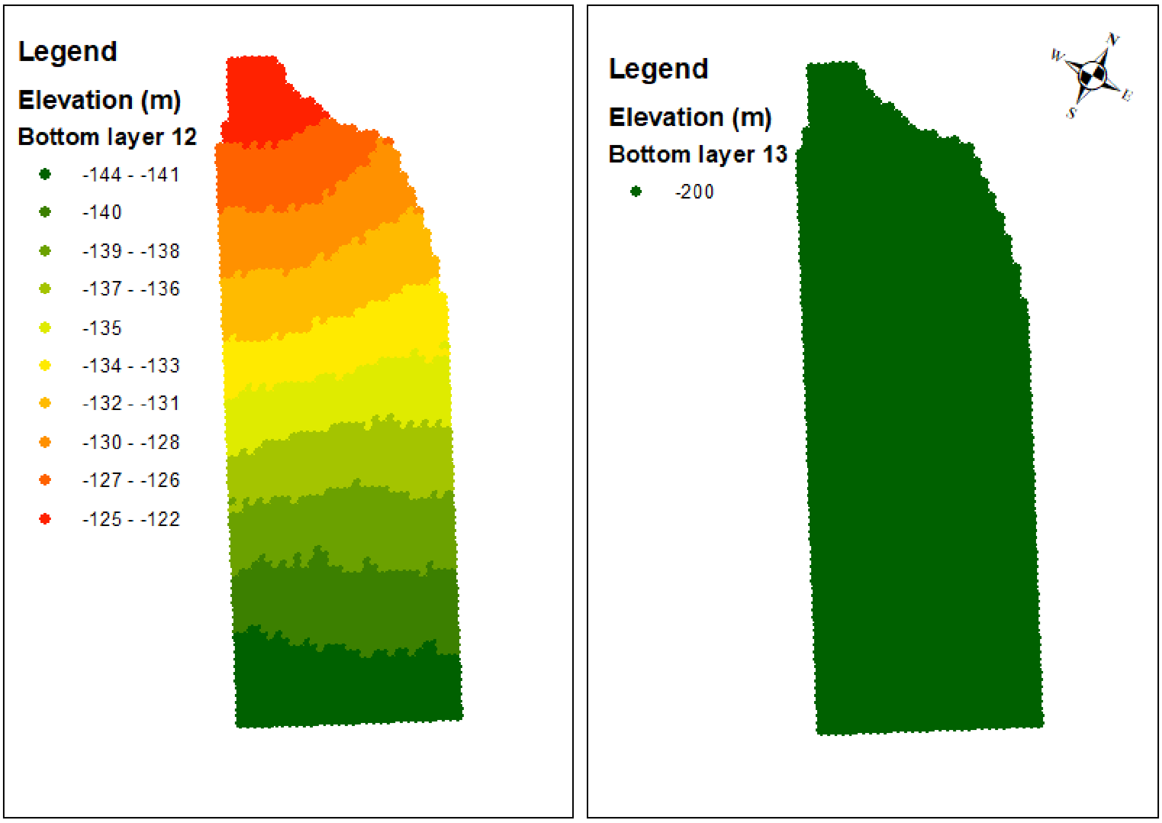

Figure C-4 Layer 12 to layer 13 elevations.

Appendix D NSMC simulations

In the initial NSMC modelling, the parameter randomisation at the start of the process sampled the original Jacobian matrix and then allowed multiplication of the calibrated hydraulic conductivity parameters by between 0.5 and 2 times. The model was then passed through two calibration cycles to optimise the results to the available observation. Because the results of the first NSMC appeared on analysis to be too constrained, the sampling and recalibration criteria for the second NSMC suite were relaxed. The NSMC runs produced a suite of calibrated models that all performed satisfactorily against observation data. The model percentile statistics can be presented in several ways. Maps can be produced that present the combined mounding percentile on a cell by cell basis, and alternatively, others can be produced that present a holistic model domain mounding response. Producing statistics on a cell by cell basis is closest to the analytical suite presented in 5.1. On a cell by cell basis, this is the most robust statistical measure of potential changes. However, this approach loses the parameters that went into producing the output, as it is likely the results would be from a combination of models. Alternatively, generating the statistics from the total model domain mound still produces a direct linkage to the actual model that produced the result. Thus, it is possible to know the parameters that produced the results and reuse the model for further scenario testing.

Figures D-1 to D-3 present the calibration fit of the simulations equivalent to the respective percentile in terms of total model domain mounding, while Table D-1 shows the calibration statistics.

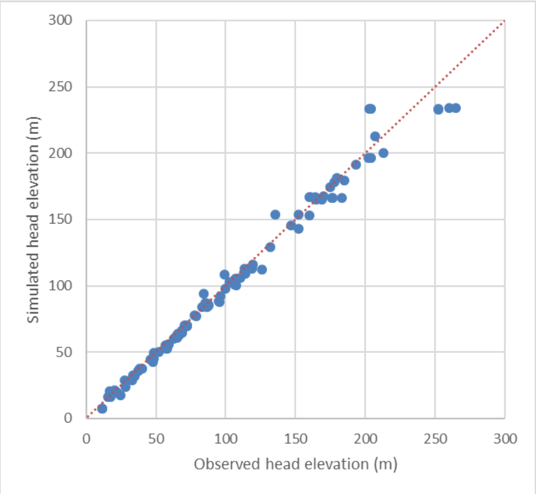

Figure D-1 5th percentile.

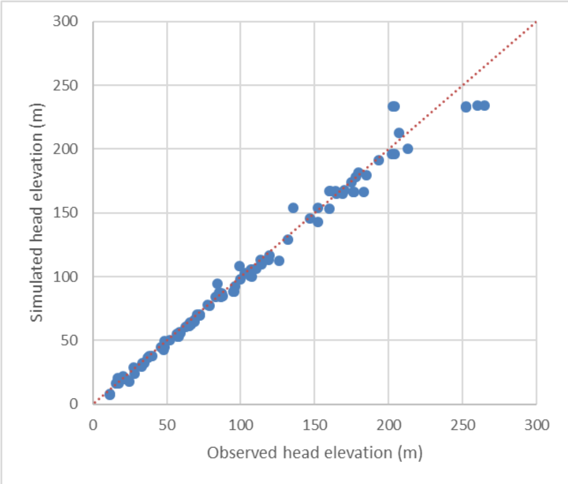

Figure D-2 50th percentile.

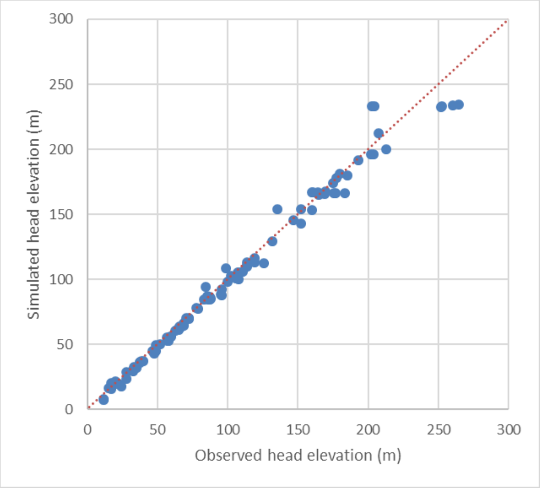

Figure D-3 95th percentile.

Table D-1 Calibration statistics

Statistical measure	*Base model (m)*	*5th percentile (m)*	*50th percentile (median) (m)*	*95th percentile (m)*
Residual mean	2.06	2.42	2.36	2.39
Residual standard deviation	9.13	7.84	7.84	7.84
Absolute residual mean	6.37	5.18	5.15	5.16
Residual sum of squares	8.58 x10^3	6.59 x10^3	6.57 x10^3	6.58 x10^3
RMSE	9.36	8.2	8.19	8.19
Minimum residual	-31.29	-30.82	-30.85	-30.72
Maximum residual	31.29	30.39	30.36	30.49
Range of observations	253.12	253.12	253.12	253.12
Scaled residual standard deviation	0.036	0.031	0.031	0.031
Scaled absolute mean	0.025	0.02	0.02	0.02
Scaled RMSE	0.037	0.032	0.032	0.032
Number of observations	98	98	98	98

NSMC Mounding

Perhaps surprisingly, the extent of mounding change between simulations was minor (Figures D-5 and D-6). This is in quite stark contrast to the 3500 random parameter simulations generated using the analytical solutions, which showed significant variance in simulated mounding. This perhaps suggests the base model may be close enough to the probable outcome. An alternative and more likely speculation is that the NSMC parameter set was too constrained in the Jacobian Matrix starting values and perhaps allowed too many calibration runs before model termination.

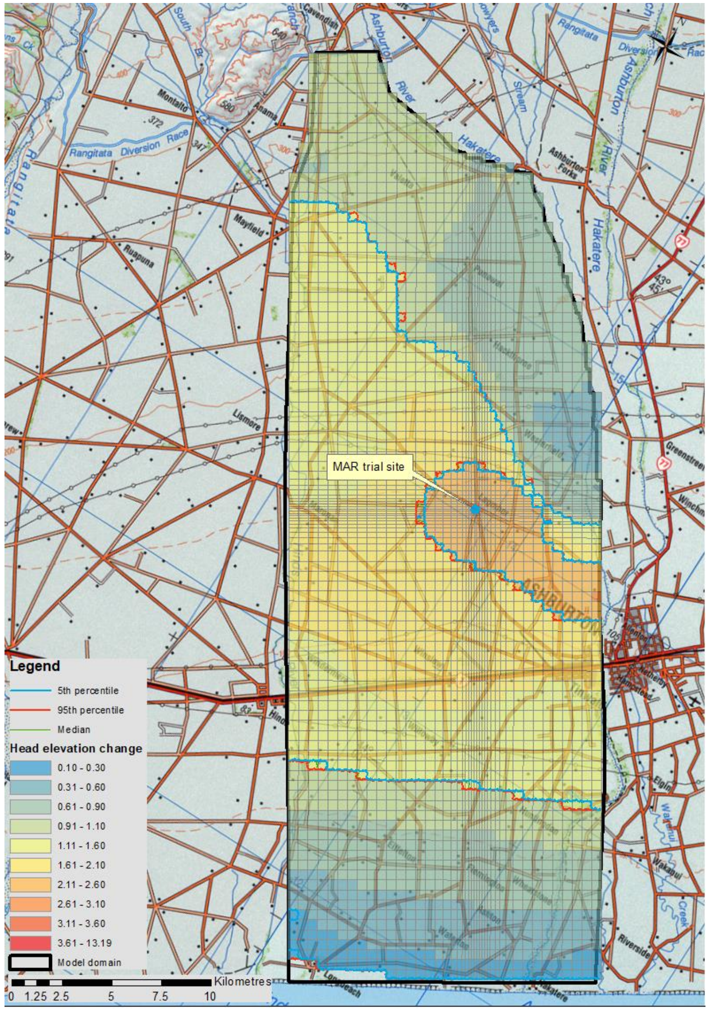

Figure D-5 Water table mounding (general head boundary).

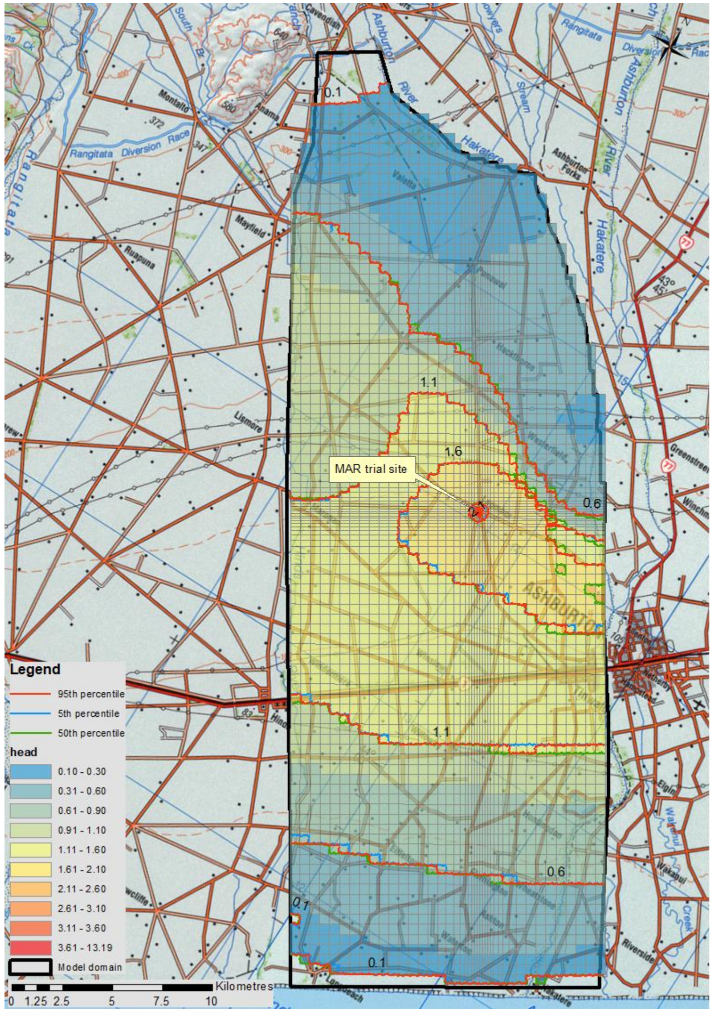

Figure D-6 Water table mounding (fixed head boundary).

Transport model results using 5th, 50th and 95th percentile flow fields

Water quality transport was added to the NSMC generated flow models using the USGS software MT3DMS. The transport models used a constant flow field; they were themselves run as transient simulations. Three flow fields representing the 5th, 50th and 95th percentile total domain mounding were used in the transport modelling. Each transport model was run across five years, with results reported after five years. As discussed in Chapter 4, because both background conditions and loads entering the groundwater system were unknown across the model domain, the MAR water was modelled as the contaminant plume itself. The recharge water was given a uniform value of 100 mg/l, with the background conditions set to zero mg/l.

The suite of simulations was generated by varying the porosity of the saturated zone uniformly across the model domain from 0.01 to 0.3. Manual calibration of porosity to the two observation wells BY20/0152 and K37/0200 suggest the porosity is between 0.01 and 0.05. The results for all the simulations were then merged and the average MAR plume calculated. Figures D-7 to D-14 show the average plume simulated by the model. Manual calibration to the observation wells suggests the water quality results lie closer to the 95th percentile. Overall, a poor fit to the observed concentration data was achieved. In general, the modelled plume moved too far to the south and avoided the observation wells, which showed a response.

Figure D-7 Average MAR plume Layer 2.

Figure D-8 Average MAR plume Layer 3.

Figure D-9 Average MAR plume Layer 4.

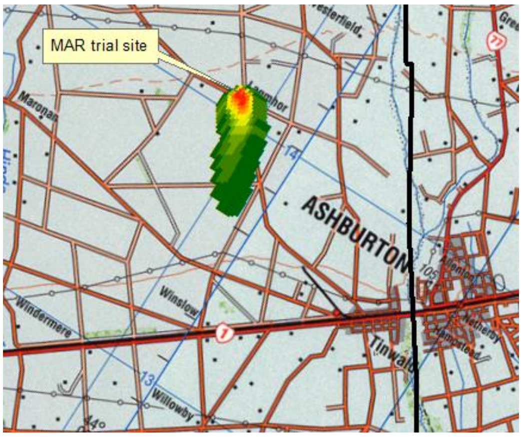

Figure D-10 Average MAR plume Layer 5.

Figure D-11 Average MAR plume Layer 6.

Figure D-12 Average MAR plume Layer 7.

Figure D-13 Average MAR plume Layer 8.

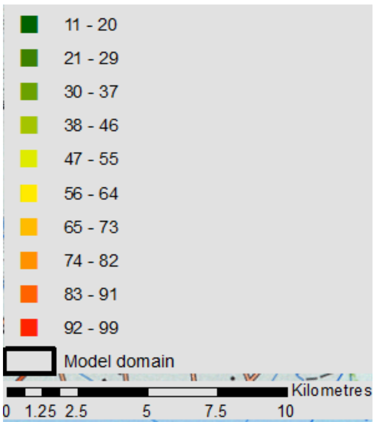

Figure D-14 Percentage of water from MAR plume.